NORMATIVE VISUAL DEVELOPMENT:
INNATE LEARNING IN THE EARLY VISUAL SYSTEM

A Dissertation

Presented to the Faculty of the Graduate School

of Cornell University

In Partial Fulfillment of the Requirements for the Degree of

Doctor of Philosophy

by

Mark Vincent Albert

February 2010

NORMATIVE VISUAL DEVELOPMENT:

INNATE LEARNING IN THE EARLY VISUAL SYSTEM

Mark Vincent Albert, Ph.D.

Cornell University 2010

Traditionally, visual development is thought to occur in two distinct stages, an innate stage which occurs before eye-opening and a learning stage in which experience shapes development. This dissertation will show how the traditional view represents a false dichotomy limiting progress in understanding visual development. Prior to eye opening there is patterned, spontaneous neural activity. Models will illustrate that spontaneous neural activity can shape development in the same way as early, natural visual experience; the same learning method can be used in both stages of development. A parsimonious view of early development is proposed - a method of 'innate learning' that prepares the visual system for later experienced-based refinement by learning from endogenous, patterned neural activity.

This dissertation presents a series of research studies which uniquely apply a normative, efficient coding approach to understand visual development prior to eye opening. The neural code in fully-developed adult primary visual cortex (V1) is presumed to be efficient. By efficiently encoding images of our natural visual environment, techniques such as sparse coding and independent components analysis (ICA) have produced linear filters which resemble experimentally measured receptive fields in V1. Simulations in this dissertation show how spontaneous patterns of activity are capable of producing these V1-like visual codes using the same learning algorithms. The spontaneous activity models created here are abstractions of known retinal, LGN, and V1 physiology; they relate mathematically to simpler theoretical

models used in the physics of critical phenomena (e.g. percolation networks) but have parameters with direct physiological interpretations. Unlike previous models of V1 development, the innate learning principle applied here will be shown to constrain the statistical structure and form of modeled spontaneous activity to more closely resemble measured spontaneous activity patterns. The monocular activity model is later extended to include binocular neural activity; binocular disparity selective cells are refined by spontaneous activity correlated between eye-specific neurons in the LGN and V1 to promote early stereopsis. The implications of this normative approach to visual development will be explored in detail, showing how this 'innate learning' approach can greatly improve our understanding of sensory development.

BIOGRAPHICAL SKETCH

Mark V. Albert graduated from Pittsburg State University in 2000 with majors in Chemistry, Mathematics, Physics, and Computer Science. After a year abroad studying neural networks as a Fulbright scholar in Vienna, Austria, he began his research career at Carnegie Mellon University. There he conducted research in primate visual neurophysiology, human problem solving using fMRI, and efficient coding applied to the visual system. It was his exposure to concepts in neural efficient coding at CMU that led him to pursue graduate school at Cornell University with David Field. At Cornell he transferred to the new field of computational biology at Cornell, while working with Field in the psychology department. He also began graduate school in the nonlinear systems NSF IGERT program. In this program, he was initially exposed to the network models which inspired much of this work. In December 2009, he will begin work as a postdoctoral research associate at Northwestern University's medical school, working with Konrad Körding. His other professional pursuits while at Cornell include being an employee at the Cornell Public Service Center where he worked with two student outreach groups to promote science enrichment in underrepresented schools. He also worked as a research consultant for Instinctiv/Predict Systems, Inc. where he designed and tested inference methods for predicting user preferences.

ACKNOWLEDGMENTS

I would like to thank David Field for his help over the years. In particular for his focus on intuition, context, and overall goals in guiding my work and the implicit attitudes toward research and teaching that have invariably become part of my motivation as well. The guidance from Jonathan Victor, Barbara Finlay, John Guckenheimer, and Shimon Edelman has also been helpful to my graduate career. I have also benefited from a number of discussions with Risto Miikkulainen in UT Austin in the past year further developing the future course of the 'innate learning' approach. I had the pleasure to work with two undergraduates on these projects – Adam Schnabel and Brandon Liu. Also, I would like to give a special thanks to John R. Anderson at CMU for demonstrating to me how a career in academia can be both interesting and rewarding. The following funding sources have made this work possible: National Geospatial-Intelligence Agency contract HM 1582-05-C-0007 to David J. Field and the NSF nonlinear systems IGERT grant and later the NSF graduate research fellowship to me.

TABLE OF CONTENTS

LIST OF FIGURES

LIST OF ABBREVIATIONS

ICA	Independent components analysis
LGN	Lateral geniculate nucleus
PCA	Principal components analysis
RGC	Retinal ganglion cells
ROI	Region of Interest
SCA	Sparse components analysis
V1	Primary visual cortex

UNIT 1: BACKGROUND

Chapter 1: Introduction to the concept of innate learning

Visual development progresses from a primarily innate, genetically-determined stage before eye-opening to an experience-driven learning after eye-opening. Much work has been done to understand these two stages in isolation; however, many of the same characteristics of visual development are present both prior to and after experience. This dissertation explores a form of learning which counterintuitively occurs before eye opening. Primary visual cortex (V1) cell responses may adapt to patterns of spontaneous activity in the early visual system in ways that resemble learning by natural visual experience. The exact nature of this adaptation is unknown, but this 'innate learning' principle has the potential to greatly simplify our understanding of sensory development. What was previously two stages of development appears to be one stage with only a change in the type of input. This principle aids not only our understanding of the nature of spontaneous neural activity but also the shared method of learning occurring both prenatally and postnatally. The computational models in this thesis will show how the known neurophysiology, as well as earlier neural models, can be explained as instantiations of this principle.

> ***The principle of innate learning:*** *spontaneous patterns of neural activity can be used to train or refine a sensory system in an analogous way to how the system can adapt based on natural experience.*

The term 'spontaneous activity' can appear ambiguous both lexicographically and as referenced in the literature. To be precise, the work presented here focuses on patterns

of endogenously-generated neural activity in the visual system prior to experience. This neural activity is spontaneous insofar as it is endogenously generated, but the activity is also highly structured – patterned, correlated bursting activity throughout the early visual system. Also, in the models presented here, spontaneous activity refers to neural activity prior to experience, although the term is more generally considered to be activity not directed by a specific stimulus or under conscious control. This is ambiguous in the adult animal, as endogenous activity in the adult can be shaped by stimuli prior to the recorded activity. This ambiguity in the term "spontaneous activity" may not be completely counterproductive; discussions in the last chapter will comment on how spontaneous activity prior to experience may relate to more mature forms of spontaneous activity like sleep, intertrial neural activity, or even neural variability during task performance.

Integrating multiple levels of analysis

There have traditionally been two levels of analysis for understanding the nature of spontaneous neural activity in early visual development, experimental and neural modeling. Some more recent experimental neurophysiology recordings have consisted of calcium imaging experiments of retinal spontaneous activity (e.g. Feller et al. 1996) and electrode penetrations of the lateral geniculate nucleus (LGN) and primary visual cortex (V1) (Chiu and Weliky 2001, Fiser et al. 2004, Schwartz et al. 1998, Weliky & Katz 1999, Yuste 1995). Although the bulk of research on spontaneous neural activity is in the retina, much of the work discussed here is based on similar patterns of activity in the LGN and V1. Early visual spontaneous activity has been characterized and manipulated pharmacologically and surgically, with a wide variety of effects discussed in the next chapter. To consolidate this wide array of experimental findings, a number of computational neural models of visual

development have been created (reviews in Erwin and Shulten 1995, Miikkulainen et al. 2005, and Swindale 1996). The most distinguishing results of these models are the ability to generate receptive fields and 2D topographic maps of V1. Primary visual cortex has a retinotopic arrangement of cells that are responsive to particular visual features, such as eye preference, orientation, spatial frequency, and direction of motion. These multiple feature maps form a patterned topography which these neural models have been able to reproduce. Unfortunately, as will be further mentioned in chapter three, there is a wide variety in the way these models accomplish this task, and very little data on spontaneous activity in the LGN and V1 for proper validation. In fact, the spontaneous activity models that have been employed range between uncorrelated noise (Linsker 1986), random lines (von der Malsburg 1973), spherically symmetric neural activity (Miller 1989), and gaussian 'blobs' (Miikkulainen 2005). Although such models can reproduce some properties of visual cortical development, the underconstrained nature of these models diminishes their ability to generalize beyond currently measured phenomena. This is particularly problematic for understanding the nature and form of spontaneous activity in the early visual system.

Fortunately, research in visual development can be guided by an additional level of interpretation. As mentioned previously, this dissertation supports the idea that the early visual system may be learning from spontaneous neural activity in a similar way to how it learns from natural experience; both spontaneous neural activity and natural experience can train/refine the visual system based on the statistical properties of the activity. This principle is not a statement of specifically how this learning is implemented, but rather how it relates to the overall strategy and goals of the early visual system – hence the word 'normative' in the thesis title. Neurophysiology is concerned with the particular details of how the visual system functions. Neural

modeling abstracts many of these details of implementation, but requires biological plausible algorithms; for example, units representing neurons, groups of neurons, or entire cortical columns interacting in simple, plausible ways. A normative, efficient coding level of analysis, on the other hand, is only concerned with algorithms which share the same overall computational objective; for example, an area in the visual system may be designed to solve an unknown computational task, and any method that similarly solves this same task – biologically plausible or not – may help us understand its role. This dissertation is focused on this normative, efficient coding level of analysis. Note that these levels of abstraction roughly correspond to Marr's (1982) levels of analysis - implementation, algorithm, and computation. The neurophysiology measurements collect evidence and permit simple insights into development, the various neural models can provide different algorithmic approaches for how development may occur, but a higher level, principled understanding will be shown to provide a clearer, more holistic understanding for why the early visual system may have evolved this way.

In chapter two, we will review what is understood about the neural responses in adult primary visual cortex. The responses of V1 neurons to the light/dark patterns of presented stimuli have been measured and later characterized using filter descriptions (e.g. 2D gabor filters for V1 simple cells) (Field & Tolhurst 1986, Jones & Palmer 1987). The efficient coding hypothesis states that the goal of this early sensory processing is to remove redundancy in sensory signals. Early interpretations of this principle focused on redundancy reduction in general (Attneave 1954, Barlow 1961) with later work analyzing the statistics of natural images explicitly (Field 1987) and suggesting specific coding strategies on natural stimuli, such as sparse coding (Field 1994). In fact, previous work has shown that these characteristic filters shapes are

consistent with the filters derived by an efficient coding for natural images (Bell & Sejnowski 1997 – independent coding, Olshausen & Field 1996 – sparse coding). It appears that through the process of evolution and adaptation, the adult visual system has become an efficient code for our natural world. However, this previous work did not address how this efficient code develops in an individual animal.

The principle of 'innate learning' provides a unique, high-level computational insight into development. The earlier modeling approaches to understanding visual development have been primarily at the algorithmic level. As will be reviewed in chapter three, this approach requires unnecessary detail for many high-level questions – details that are often presupposed in ways that cannot be easily verified by experimental measurements. By using an efficient coding level of analysis, we are able to address the nature of spontaneous activity with only the assumption that the visual system learns using our specified measure of efficiency. This approach is not blind to the fact that this type of learning must be instantiated in the brain; the assumption is simply that we are using a level of abstraction more directly related to the goals of the visual system. In the three-level interpretation of the visual system, a normative, efficient coding approach is not in competition with other levels of analysis, but rather it complements, constrains, and further informs neural models and experimental neurophysiology.

What has been discovered by applying the principle of innate learning

The innate learning principle is proposed as a parsimonious understanding of early visual development. This dissertation will show that learning algorithms used in computational vision and efficient coding models of visual neuroscience apply not only to adult vision, but also to development. Also, the statistical structure found in

natural scenes - structure that promotes the formation of a V1-like code - is present in spontaneous (endogenously-generated) patterns activity. That is not to say that all statistical properties of spontaneous activity are similar to activity from natural scenes (that is clearly not the case) but the low-level statistics which may help form a functional visual code can be shared between endogenous and naturally evoked activity. These approaches may appear to complicate modeling efforts by relating domains which have appeared independent, however the traditional research boundaries did not provide the means to concisely answer a number of questions. Why does spontaneous activity in the LGN and V1 appear as we observe in measurements? What purpose can mechanisms used to adapt to natural scenes, like visual critical periods, serve prior to experience? What activity patterns can train V1, assuming innate learning applies? Although this dissertation introduces additional analytical tools to understand development, the overall result is a more parsimonious understanding of spontaneous activity and learning in visual development.

A monocular model demonstrating the innate learning principle is described in chapter four. The spontaneous activity model is an abstraction of hypothesized LGN and V1 developmental activity, and resembles percolation networks used in applied math and physics. The activity model is abstract enough to have only three parameters, with each parameter having a direct neurophysiological interpretation. Given the right choice of parameters, not only do the simulated activity patterns resemble physiological patterns, but learned filters also appeared V1-like (e.g. 2D gabor filters). An exploration of the spontaneous activity parameter space shows how the filter properties can be altered, including how pharmacological manipulations of this activity could affect the produced visual code. Most importantly, the initial model shows how to validate and apply this new high-level statistical approach to

development. Previous neurophysiological recordings demonstrated correlated activity between nearby neurons, but we know from work on efficient coding applied to the adult visual system that simple correlated activity is not enough to train the visual system (Field 1987). Higher-order statistical structure, addressed in chapter three, is not only a critical component in the processing of visual scenes after eye opening, but may also be a necessary component in spontaneous patterns of activity. This insight can help to create better analytical tools, which evaluate more than just pairwise correlations (second-order statistics), to improve experimental approaches to visual learning and development.

Although the goal of the monocular model was to simulate LGN/V1 spontaneous activity, much of the modeling effort was based on work in retinal waves. This is because the vast majority of experimental work - and the only detailed computational neural models of spontaneous activity - have been in the retina. Chapter five explores a model of LGN activity which makes the distinction from retinal activity clear by including binocularity. For binocular neurons in V1 to form or refine through spontaneous activity coordinated activity has to occur between the eye-layers, as has been measured experimentally in the LGN (Weliky & Katz 1999). The binocular neural activity was encoded in a similar way to the monocular model to produce physiological filters. Activity with only partial coordination between eye-layers was able to form an adult-like distribution of disparity-selective cells; a distribution found in newborn primate V1 (Maruko et al. 2008). As features such as binocularity are added to the spontaneous activity model, additional insights into visual development are possible. For example, binocularity produces relative horizontal displacements between the images of the eye; however, the current binocular spontaneous activity model generates patterns which have equal horizontal and vertical displacements,

which produces equally horizontal and vertical displacements in the receptive fields. It has been experimentally shown that horizontal displacements are significantly more common in the adult, matching a bias that one would expect based on natural binocular experience (Cumming 2002). Are these vertical disparities equally present in activity or resulting filters present at birth? Efficient coding models can suggest experiments like this one as they frame our understanding for further experimental testing.

The binocular model also lays the foundation for an additional evaluation of innate learning approaches. Previous models express results of learning in terms of physiology (receptive fields or V1 topographic maps) or the abstract concept of efficient coding (e.g. statistical independence, log likelihood for a given data set). However, the binocular model also allows for the calculation of binocular disparity – an aspect of depth perception – which can be optimized and tested. The activity of a set of receptive fields from the binocular model can be combined to determine the disparity from a stereo pair of images. A given set of binocular filters can be tested by analyzing their ability to find the true depth map from a stereogram. In the innate learning paradigm, spontaneous activity parameters can be chosen to maximize the system's ability to make depth discriminations before training with natural stereo input. This maximization provides a clear opportunity to demonstrate the evolutionary advantage of the innate learning paradigm. In effect, animals which are more quickly able to develop a behaviorally significant perceptual ability, like depth discrimination, are more likely to survive and reproduce. The initial stages of this perceptual interpretation are shown and discussed in the later part of chapter five.

Immediate implications and directions for normative visual development

There are a number of additional insights and future directions made possible by applying the innate learning principle – the improvement in current modeling approaches to V1 development, the use of evolutionary algorithms to understand species differences and advantages, and using V1 development as an example for the broader application of the innate learning principle beyond vision and even development. It was mentioned that the neural models are not well constrained by what is known about visual neurophysiology during development. This has lead to a wide variety of V1 development models. The principle of innate learning, if it is to be followed, provides a high-level constraint on the types of acceptable neural models – ones which are able to form appropriate visual codes both on spontaneous activity as well as natural visual images. Only a few of the dozens of models have been tested in this way (e.g. Burger and Lang 1999; Bednar 2002; Bednar and Miikkulainen 1998, 2004). These neural models, to be discussed later, require richer spontaneous activity models than ones which don't consider an innate learning approach. What they are lacking is understanding the extent to which the statistics of the neural activity or the constraints of the chosen neural architecture dictate the formation of receptive fields. Efficient coding approaches, as the name suggests, rely more on the statistical properties of the neural activity to derive an encoding, thus forcing a more extensive activity model rather than a more elaborate learning architecture.

These higher levels of analysis – coding and neural modeling – do not need to be approached in isolation. The same statistical evaluations of filters from efficient coding (e.g. measuring sparseness/independence on natural images) can also be applied to filters from neural models. Equivalently, constraints (e.g. local interactions) can be added to efficient coding models to produce neural modeling-like results such

as topography (Hyvarinen & Hoyer 2001). Links between these levels of analysis in development may also relate to recent physiologically-plausible instantiations of sparse/independent coding (Falconbridge et al. 2006, Garrigues & Olshausen 2008, Rozell et al. 2008). This dissertation represents one area - development - where these two levels of analysis converge, and can significantly impact the direction of research in both levels.

As will be shown later, the simplicity of abstraction by efficient coding helps in easily extending the current modeling approach. The binocular model of chapter five uses the same efficient coding procedure used in the monocular model to derive receptive fields. Suggested future work includes a spatiotemporal spontaneous activity model to test assumptions about processing of visual motion in early development, possibly using the same efficient coding procedure but with spatiotemporal data. The innate learning principle may not only better constrain current models of visual development, but also provide a simpler framework for addressing additional features in the neurophysiology of spontaneous activity.

Chapter five addresses an additional direction of this work. The production of an efficient visual code is not only an abstract goal but has clear perceptual implications, such as early binocular depth discrimination – an important survival advantage. Also, there are distinct differences in visual development between species that can only be addressed by noting the different evolutionary constraints – for example, parental care vs. precocial young. There is a tradeoff between the ability to learn flexibly and to learn quickly – analogous to the bias/variance tradeoff in machine learning. The poor monocular properties of neural receptive fields in newborn primates and kittens (e.g low spatial frequency - Chino et al. 1997, Freeman & Ohzawa 1992) may be framed as

an evolutionary choice to optimize under the additional assumption of a limited amount of natural experience. This is distinctly different from the constraints placed on innate learning in deer, which have to be capable of visual perception quickly after birth. The implications of the evolutionary approach to understanding this developmental difference will be discussed in chapter seven.

Broad implications

The final chapter was included to address broad implications of this normative approach to visual development. Though only indirectly implicated by the current work, the possible impact on other areas of work should not be overlooked. For one example, the eventual goals of this work are not only model formation and validation, but also possible application. Because these V1 models simulate the visual processing of a wide variety of animals including humans, a number of key insights into visual processing are possible. First, this work has implications for understanding human visual impairments. Amblyopia, defined as poor vision due to abnormal visual experience early in life, affects 1-4% of the population (2-4% in Ciuffreda et al. 1992, 1-3% in Webber & Wood 2005); this condition can be present at birth with no demonstrable physical impairments. The work in this dissertation models prenatal neural activity and the manner in which it prepares the early visual system before the eyes open. This is the time at which diseases such as congenital amblyopia can develop. Current neural models cannot adequately address the normal properties of spontaneous neural activity; however, the efficient coding methods used in this dissertation require a specificity in the form of spontaneous activity which may permit a better understanding of not only normal visual function, but also abnormal development, as may be the case in some forms of congenital amblyopia.

This proposal is one step in applying a principle which may go well beyond V1 development. There are later visual areas which also must be able to function at or near birth. There are even particular biases, such as the infant orientation to faces, which may be influenced by spontaneous activity later in the visual system; this will be covered in chapter eight. In addition to such later visual areas, other sensory systems may also operate under the same principle. The same efficient coding algorithms that explained the receptive fields of V1 cells (Olshausen & Field 1996, Bell & Sejnowski 1997) also apply equally well to explaining neural responses in the auditory nerve (Lewicki 2002). Spontaneous neural activity in development occurs in the auditory and motor systems, and given the high-level coding similarities, these systems may be able to benefit from a comprehensive innate learning approach as well.

This work leverages the similarity in the adaptive learning algorithm before and after external experience, however, another link between these two stages will be addressed in chapter eight. Although the nature of spontaneous activity after initial experience changes dramatically, endogenous neural activity persists throughout adulthood. There is also recent work linking adult spontaneous activity to probabilistic inference based on evoked, natural activity (Berkes et al. 2009, Hoyer & Hyvarinen 2003). The activity patterns in these models may be related to the average evoked activity. In bayesian statistical terms, the spontaneous activity modeled here can be considered an evidence-based prior for learning a neural code, one which was instilled by the evolutionary history of a species. Although this thesis focuses on the learning algorithm being shared between two stages of development, the link between the role of spontaneous neural activity prior to experience and the role of adult spontaneous activity will be briefly discussed.

The convergence of multiple levels of analysis (neural modeling and efficient coding) and stages of development (prior to and after initial experience) offer a number of additional insights and a parsimonious interpretation of multiple phenomena. Binocularity, depth discrimination, spatiotemporal modeling, and other extensions help to validate the innate learning principle. As will be explored, these implications go beyond simply understanding V1 development. The innate learning principle discussed in the following pages may also relate to the development of higher-level vision, other sensory systems, and even adult visual processing. Before this work and its implications are discussed, the next two chapters will review the necessary physiology and computational methods.

Chapter 2: The current role of spontaneous activity in the visual system

The modeling results to be presented involve many known aspects of visual neurophysiology. The spontaneous activity model is derived from retinal wave physiology with binocular interactions in the second version that are strictly in the LGN and V1. The first recordings of spontaneous, patterned bursts of activity in development were in the retina (Galli & Maffei 1988); the retina remains the most thoroughly studied source of visual spontaneous activity. Although the activity models used in this thesis are meant to resemble LGN/V1 physiology, a review of the work on LGN and V1 spontaneous activity will show why a certain reliance on retinal wave physiology is necessary. Additionally, the experimentally measured effects of spontaneous activity on V1 will be reviewed. The simple role of retinal wave physiology for topographic refinement in the LGN will be contrasted with the role of later spontaneous activity on the multifaceted development of V1 neural selectivity. Although ocular dominance (OD), orientation (OR), and direction selectivity (DR) are all part of the adult V1 topography in animals like the ferret, cat, and monkey, we will see how these features develop in separate stages. In the ferret some have claimed that OD columns form without spontaneous activity (Crowley & Katz 2000), others found that spontaneous activity is necessary to establish as well as maintain OR selectivity for the majority of V1 cells (Chapman & Stryker 1993, White et al. 2001), while conversely, DR selectivity requires visual experience after the eyes open (Li et al. 2006). This physiological review will demonstrate that the source of V1-relevant spontaneous activity is more complicated than retinal waves (e.g. LGN and V1 interaction), and the specific role of innate learning is situated in a developmental timeline which dictates what statistical features are critical, and which may not be present in the activity.

Visual system introduction

This section is provided to orient these results and provide context for those unfamiliar with the characterization of neurons and pathways in the early visual system. The visual system has been studied experimentally in a number of mammalian species, including human, monkey, cat, and ferret. For a number of reasons, many of the important results are only measured in a subset of these species, but in general they apply to other species as well. This dissertation attempts to generalize to the human visual system, but relies on data from other animals - mostly ferret and cat in this chapter - where human data is not available. This work additionally focuses on the thalamocortical pathway, where visual processing occurs first in the retina, then the lateral geniculate nucleus (LGN) in the thalamus, and finally reaches primary visual cortex (V1). In primates and many other higher-order mammals this is the primary visual pathway, and in humans the loss of this pathways leads to an inability to consciously perceive visual information. An important alternate route for visual information is the tectal pathway, which begins in the retina and eventually reaches the superior colliculus or pulvinar. The tectal pathway in primates is used for more primitive visuomotor functions, like directing attention and orienting eye-gaze. Here we focus on the thalamocortical pathway as it is the primary route of visual information in humans, and is the most extensively studied visual developmental pathway in mammals.

Retina: Visual information is first transduced from light to electrical activity in the photoreceptors of the eye. In humans this is with two classic types of photoreceptors – rods and cones. Rods are capable of transduction at lower light levels, while the three types of cones are sensitive to different spectral frequencies of light which allows for

color perception. Without loss of generality this dissertation will be limited to luminance only. The transduction of light intensity to electrochemical signals by these photoreceptors is often approximated as a log transformation, although more sophisticated models exist (van Hateren & Snippe 2006). From the photoreceptors, the processing of visual information occurs into two distinct layers – the outer and inner plexiform layers of the retina. The outer layer consists of horizontal cells, which collect and modulate visual information through horizontal connections, and bipolar cells, which pass this information to the inner plexiform layer. In the inner layer, which generates the spontaneous activity in the retina, amacrine cells integrate and modulate the information passed from the bipolar cells. There is wide variability in amacrine cells types, with many integrating information horizontally in the retina. One type of amacrine cell, the cholinergic starburst cell, is responsible for the generation and propagation of at least one stage of retinal wave activity (Zhou 1998) and will be discussed later. The final destination of this processing is the retinal ganglion cell; the axonal projections of these cells make up the optic nerve. In all, this processing accounts for an approximately 100:1 reduction in the number of photoreceptors in the retina to the number of axons in the optic nerve.

The type of visual processing performed by the retina can be understood by noting the area in visual space which affects the response of a neuron – the neuron's receptive field. Retinal ganglion cells perform contrast processing using a center-surround receptive field. When measuring from a single retinal ganglion cell (or equivalently, an axon of the optic nerve), the neuron fires when there is a difference in light intensity between a particular location in retinotopic coordinates and the surrounding area. ON-center cells respond to increased light intensity in the center and decreased intensity in the surround, while OFF-center retinal ganglion cells have the opposite

preference. The receptive fields are often displayed as concentric rings of differing light intensity preferences. Some have referred to this type of processing as a compressive code (Atick and Redlich 1992) which minimizes the number of axons necessary in the optic nerve, although later work suggests that this is an oversimplification (Graham et al. 2006). Also, it should be noted that even at this stage, the simple receptive field description is inadequate as these receptive fields can change dramatically (e.g. losing their surround suppression in low light).

Lateral Geniculate Nucleus (LGN): On the way to the LGN, a portion of the axons in the optic nerve cross at the optic chiasm. After this crossing, the left and right tracts no longer represent the left and right eyes, but rather the contralateral visual field. This is evident in the organization of the LGN in many animals where, for example, the right LGN processes images in the left visual field. The LGN in humans is composed of six layers. These layers can be summarized by their response properties and connectivity. Layers 1, 4, and 6 represent inputs from the contralateral eye, while layers 2, 3, and 5 are from the ipsilateral eye. Also, layers 1-2 make up the magnocellular pathway while layers 3-6 are parvocellular. The main distinction here is that magnocellular cells have higher temporal precision, less spatial resolution, and almost no color discrimination, while parvocellular cells require more time to respond, are more spatially precise, and are more color selective. Within each of these six layers, there is a separation of ON and OFF center cells. All layers show a corresponding retinotopic arrangement. Although the LGN is known to modulate activity from the retinal ganglion cells, the response properties of LGN neurons match the responses of the incoming retinal ganglion cells with high fidelity, especially in attentive/aroused states (Ozaki & Kaplan 2006). For this reason, the LGN has been

referred to, in a sweeping oversimplification, as a 'relay' for retinal signals reaching cortex.

Primary Visual Cortex (V1): Axons from LGN neurons project to primary visual cortex. The response properties of neurons in visual cortex are more complicated, and the spatial arrangement of these neurons goes beyond simple retinotopy and layer segregation. V1 cells can be classified by their response sensitivity to sine wave gratings – smoothly varying patterns of light and dark stripes. A particular V1 neuron is sensitive to the orientation, spatial frequency, and particular location of the grating. V1 simple cells are additionally sensitive to a particular phase of sinusoidal grating, whereas complex cells are insensitive to the phase of a presented grating. In this work, we will be primarily concerned with simple cells in V1 as they are more directly related to the linear receptive field models discussed here. V1 simple cells can be described by a receptive field profile with elongated alternating regions of light and dark preference parallel to each other. These regions form a characteristic pattern, first described qualitatively as gabor wavelets (Daugman 1980, Marcelja 1980), although there are other mathematical descriptions, such as 2D hermite functions, are also possible (Victor et al. 2005). Referring to the initial formalism, the shape of receptive field is often referred to as a gabor function, a mathematical function which fits the stimulus-response properties of the neurons reasonably well (1D fit from Field & Tolhaurst 1986, 2D fit from Jones & Palmer 1987). 2D gabor functions are planar sine waves of a characteristic orientation and spatial frequency in a gaussian envelope at a particular location in visual space. As can be seen by such a description, neural receptive fields can directly characterize the particular response selectivities of the neurons they refer to. It should be noted, however, that the linear filter model, which includes gabor filter descriptions of cell response, is a crude description of a much

more complex neural response. Neural responses are clearly nonlinear, as they cannot fire negatively and have a maximum response, among many other nonlinearities (reviewed in Albrecht et al. 2003); however, a linear model will be assumed as it is much simpler to describe mathematically and covers many of the basic properties of V1 simple cell neural response. As we will see in the next chapter, the functional/computational purpose for the particular shape of these receptive fields can alternately be explained by principles of efficient coding.

In addition to more complicated receptive field properties, the spatial arrangement of V1 cells is more complex. Many of the V1 cell selectivities (location, spatial frequency, orientation, ocular dominance, direction of motion) are arranged in cortex to have neurons with similar selectivities located near each other. In order for this to be accomplished on a 2D cortical map, various overlapping and interleaving patterns emerge. Although corresponding retinotopic locations in the LGN project to corresponding locations in V1, there are minor adjustments due to the requirements of these overlapping patterns. For example, within the same corresponding visual location, LGN cells from the two different eyes project differentially to two neighboring, but separate stripes on the surface of cortex – known as ocular dominance columns. Within these columns, neurons response preferentially to stimuli in one eye over the other. Overlapping on this map, orientation selectivity varies smoothly over the cortex, and a direction selectivity map overlays on the pattern of orientation columns. In some locations, known as blobs due to characteristic cytochrome oxidase staining, there are low spatial frequency cells, while the interblobs have higher spatial frequency selective cells. All of these various arrangements of cells are intertwined in a roughly retinotopic map of visual space. As will be explained in the computational maps section of the next chapter, this spatial

arrangement of cells is likely to be a developmental mechanism associated with the requirements of lateral connectivity between similar cells and is critical for instantiating some of the processing requirements of vision.

Retinal waves

Baseline neural firing rates, a form of spontaneous activity, have been observed ever since the earliest *in vivo* neural recordings; however, the critical experimental insight in perinatal visual development was discovering the presence of patterned/correlated neural activity. Spontaneous, correlated bursts of neural activity during visual development were first measured in the retina (Galli & Maffei 1988, Maffei & Galli-Resta 1990 - reviewed in Wong 1999). This type of coordinated activity occurs before the eyes first open and has been found to occur in rats (Galli & Maffei 1988), macaques (Warland et al. 2006), ferrets (Feller et al. 1996), rabbits (Zhou 1998), and kittens (Meister et al. 1991) as well as non-mammalian species including chicks (Catsicas et al. 1998, Wong 1998) and turtles (Sernagor & Grzywacz 1993). Calcium imaging experiments of intact retinae have shown the qualitative nature of these retinal waves. The activity spreads as amorphous patterns of bursting neural activity with relatively slow dynamics. For example, in the ferret retina, the average time between retinal waves in a given location is a median 115 seconds (Feller et al. 1996), a typical ganglion cell's burst of activity lasts roughly one second, and the average wave propagation speed is 177 micrometers/second (Feller et al. 1997). The spontaneous activity is known to originate in the amacrine cell layer, with starburst amacrine cells bursting independently and then recruiting each other into waves of activity (Zhou 1998). The ganglion cell layer essentially acts as a low-pass filter (Butts et al. 1999), passing this activity with minor modification to the LGN.

Combining a number of research results, these waves have been characterized by neural models that reproduce aspects of the activity patterns such as wave shape, speed, size, and duration using physiological details like dendritic field size and measured cell spacing (Burgi & Gryzwacz 1994, Butts et al. 1999, Godfrey & Swindale 2007). The two most recent retinal wave models simulate the independently bursting, refractory, and recruiting behavior of the amacrine cell layer. The Butts model additionally produced retinal ganglion cell responses while the Godfrey and Swindale model simply smoothed the amacrine cell activity. In both models, amacrine cells could be in one of three states: active, refractory, or recruitable. The Butts model simulated ferret retinal wave activity while the Godfrey and Swindale model simulated retinal wave activity in ferrets, rabbits, mice, chicks, and turtles by changing only a few parameters. In the Butts model, there were two free parameters unconstrained by direct neurophysiological measurements – the threshold number of neighboring amacrine cells needed to make a recruitable amacrine cell active and the rate of independent initiation of a burst. What was found by the model is that the propagation speed of a wave was directly related to one overall parameter – the fraction of recruitable amacrine cells; this last parameter will be discussed in relation to the activity model used in chapter four. These two most recent neurophysiological models are fairly consistent in their interpretation of retinal wave activity and provide the basis for much of the later inferences to endogenous activity in other parts of the visual system.

While research efforts were continuing in characterizing these waves, the effects of these waves were also being explored. Experiments have manipulated these waves by abolishing them, over-stimulating them, or otherwise altering their properties and have shown how they are necessary for proper development (Cang et al. 2005, Chapman et

al. 1986, Chapman and Godecke 2000, Lowel and Singer 1992, Stryker and Strickland 1984, Weliky and Katz 1997). Molecular guidance cues are capable of directing axons to their approximate retinotopic locations in the LGN, but waves are necessary for the refinement of these connections (Shatz and Stryker 1998). These retinal waves may also have a role in the retinotectal visual pathway by refining the retinotopic map to superior colliculus (Chandrasekaran 2005). In relation to V1, the retina is not directly connected, so the indirect effect of retinal wave activity on V1 will be discussed through its connection to the LGN in the following section.

LGN and V1 spontaneous activity

In the visual system, both the LGN and V1 have patterned, spontaneous activity during development (Chiu and Weliky 2001, Fiser et al. 2004, Schwartz et al. 1998, Weliky & Katz 1999, Yuste 1995). In comparison to retinal spontaneous activity, later visual spontaneous activity is not as thoroughly characterized or experimentally manipulated. One reason for this is the ease of a retinal preparation and imaging versus electrode recordings; also interpretation is more difficult as the later visual stages are less isolated. Spontaneous activity in LGN, for example, is a dynamic process affected by both retinal waves as well as feedback from V1. Although there is no full characterization of LGN/V1 activity, the proposal is that this later visual spontaneous activity is closely related to the concept of innate learning.

Retinal waves, where the majority of experimentation on visual endogenous activity has focused, cannot be the sole source of innate learning patterns for V1. One immediate objection is that the waves are not binocular. Days after birth, primates have an adult-like ocular dominance distribution (Chino et al. 1997) and have an adult-like distribution of binocular disparities (Maruko et al. 2008). Retinal waves are

not correlated between the eyes to promote or maintain such selectivity. Another criticism is that in some animals, like the ferret, retinal wave activity only minimally overlaps in time with the onset of V1 activity. Retinal waves end between postnatal day 24 and 30 whereas V1 cells become responsive on approximately day 23 (reviewed in Wong 1999). Also, the retinal waves of the ON and OFF-center cells are more independent at later stages (Wong & Oakley 1996), which is helpful for segregating ON and OFF layers in the LGN, but the responses of these cell classes are intimately related in V1 neurons. Such concerns have turned many researchers away from considering the possibility of an innate learning role for spontaneous activity.

Although retinal spontaneous activity cannot fill this role as an innate learning training pattern, it appears that LGN spontaneous activity can. The endogenous activity in the LGN also overlaps in time with V1 responsiveness (Wong 1999). The activity is correlated between eye-specific layers, as well as ON and OFF-center cell layers (Weliky and Katz 1999). Finally, the evidence from electrode arrays shows that the activity appears as a wavefront moving across these layers. LGN endogenous activity appears better suited as a potential source of innate training patterns.

The retina is a relatively isolated structure, making characterization of retinal waves simpler. LGN and V1 activity, on the other hand, are complicated by the effects of extrinsic connections as well as intrinsic circuitry. Weliky and Katz (1999) systematically measured the effect of manipulating the connections between visual areas on spontaneous activity in the LGN. First, after removal of retinal inputs by severing the optic nerve, LGN activity was temporarily abolished. Over the course of approximately six hours, LGN activity gradually reappeared and stabilized, but was slightly altered with burst frequency nearly doubling. More importantly, they

measured the effect of V1 feedback on LGN activity. Correlated activity between LGN eye-specific layers was abolished when V1 feedback was removed by aspirating ipsilateral visual cortex. Although the consequence of optic nerve transection and subsequent changes in burst frequency on the innate learning role of LGN activity is not clear, certainly uncorrelated activity between eye layers has clear implications. Neurons in primary visual cortex are predominantly binocular, and binocular correlations are important in any potential training pattern. For this reason, this dissertation will make the assertion that combined LGN/V1 activity patterns drive visual development in an innate learning paradigm – though for simplicity, this combined activity may simply be referred to as LGN activity at times.

It should be noted that V1 also has intrinsically generated activity. Chiu and Weliky (2001) found that V1 spontaneous activity after postnatal day 22 in ferrets exhibited long-range correlated neural activity. The endogenous activity was found to fall off over distance, but exhibited significant long-range peaks in correlation at a distance of approximately a millimeter in cortex. These peaks in correlation are believed to be caused by either the segregated ocular dominance inputs from the LGN or patchy horizontal connections found within V1. Chiu and Weliky temporarily abolished LGN activity using binocular enucleation, as mentioned previously. The patchy correlated activity remained in the intrinsic V1 activity, showing at least some independence from LGN activity. Also, this work showed that macroburst frequency, both before and long after binocular enucleation, was matched between LGN and V1, showing that the patterns of activity in these two areas are intimately connected.

V1 feature maturation in relation to spontaneous activity

The role of spontaneous activity proposed in this dissertation is a bridging point for

V1 development between molecular guidance cues and refinement based on visual experience. The review of V1 maturation in this section will focus on the ferret and cat (the most studied animals in V1 development) and cover three different stages of V1 development. It has been claimed that V1 ocular dominance columns can form without the influence of spontaneous activity (Crowley and Katz 2000), but neural activity is required for maintenance; for most V1 cells, orientation selectivity requires spontaneous activity for formation as well as maintenance (Chapman and Stryker 1993); and direction selective columns do not form without visual experience (Li et al. 2006). We will review to what extent spontaneous activity can have an instructive role in formation or maintenance, or even no role at all, depending on the particular feature of interest.

Ocular Dominance (OD) columns: Ocular dominance, the preferential response to one eye over the other, is associated also with a predominance of innervation from the LGN into layer IV of V1 from the dominant eye. Primate work (Horton & Hocking 1996) showed that these patchy connections - functionally corresponding to ocular dominant neural responses - occur prior to visual experience. Equivalently, Crair et al. (1998, 2001) demonstrated that kittens deprived of visual experience by closed eyelids still develop OD bands at two weeks after birth (eyes normally open after one postnatal week). These results suggest that visual experience is not necessary for OD band formation. To demonstrate the role of spontaneous activity, Crowley and Katz (2000) binocularly enucleated ferrets before LGN inputs entered layer IV of V1. Anterograde tracing of LGN injections and retrograde labeling in V1 revealed patchy patterns of the same periodicity as control animals. Additionally, in normal ferrets, segregated patches of geniculocortal axons were seen as early as P16 to P18 – less than a week after the LGN innervates V1. Crowley and Katz go on to state that

molecular guidance cues must be the driving force in OD column formation. However, they did not experimentally address the possible role of LGN and V1 spontaneous activity in OD column formation. Although the initial mechanism for OD column formation remains unclear experimentally, visual experience is necessary to properly maintain these columns. For example, dark-reared cats have ocular dominance columns which are significantly less distinct than controls (Stryker & Harris 1986), and an imbalance of activity caused by monocular lid closure results in shrinkage of that eye's ocular dominance columns and an expansion for the other eye. Although visual experience may not be necessary for initial formation of OD columns, it is necessary for maintenance. Additionally, retinal wave activity was shown to not be necessary for formation, but the precise role of LGN/V1 spontaneous activity remains unclear, with some researchers suggesting initial formation through molecular guidance cues alone.

Orientation (OR) columns: Orientation column formation also does not require visual experience, but is more sensitive to changes in natural vision after birth. One-fourth of recorded V1 neurons in neonatal kittens prior to visual experience, as well as dark-reared kittens, show some degree of orientation selectivity (Blakemore & Van Sluyters 1975). Optical imaging of visually deprived kittens by binocular lid-suture revealed a normal topographic layout of orientation preference (Crair et al. 1998, Goedecke et al. 1997). In fact, orientation selectivity continues to improve independent of the presence or absence of visual experience up to postnatal day 18 (Fregnac & Imbert 1984). However, prolonged dark-rearing causes the orientation maps to deteriorate and eventually disappear (White et al. 2001). As with OD columns, it appears that OR columns do not require visual experience for formation, but require it for maintenance. Also, similar to the effect of activity imbalance in OD,

exposure to particular orientations early in life creates an abundance of cells tuned to that particular orientation in V1. Sengpiel et al. (1999) demonstrated that approximately twice as much cortical area in V1 was dedicated to the orientation that kittens were raised in compared to the orthogonal orientation, however all orientations were still represented in V1.

In contrast to OD column formation, however, OR columns eventually disappear without proper visual experience through extended dark rearing (White et al. 2001), showing a greater sensitivity to manipulations of visual experience. Weliky & Katz (1997) manipulated the spontaneous activity by enucleating one eye and placing a wire cuff around the optic nerve of the other eye. They proceeded to induce synchronous electrical bursts. The resulting orientation maps appeared normal, although orientation selectivity for individual neurons appeared significantly lower demonstrating an effect of alterations to spontaneous activity. However, this manipulation also did not address the role of later spontaneous activity in the LGN or V1. Chapman and Stryker (1993) injected visual cortex with the sodium channel blocker TTX in order to stop visual activity. This had the effect of completely blocking orientation column maturation. Therefore, similar to ocular dominance column formation, orientation selectivity and orientation column formation may appear in a crude form prior to activity, but require spontaneous neural activity to mature normally.

Direction of Motion (DR) columns: Neurons in V1 are also selective to a particular direction of motion, typically orthogonal to the preferred orientation, with DR maps present in the adult. According to experiments by Li et al. (2006), direction selectivity in ferrets appears to form only after the eyes open. Using drifting gratings they found,

unsurprisingly, that well-defined orientation maps were present before eye opening. DR maps, however, did not form until the first week of visual experience and only reached maturity in the second week after eye opening. Additionally, they showed that dark-rearing that lead to the expected increased presence and then gradual decline of orientation selectivity over ensuing weeks never lead to the formation of direction selective cells. Direction maps only formed when dark-rearing ended within the first week after eye-opening, whereas the timing of visual experience required to maintain the orientation map was significantly longer. These results suggest that direction selectivity requires visual experience and also has a much faster critical period. This shorter critical period agrees with early experiments of kittens raised in rotating drums, where reversing the direction of the drum 5 weeks after birth still led to a preponderance of cells sensitive to the initial direction of motion (Daw & Wyatt 1976). In contrast, reverse lid suture at 5 weeks leads to a preponderance of cells more sensitive to stimuli in the recently opened eye (Berman & Daw 1977). In any case, it appears that the topographic organization of direction selectivity requires visual experience for initial formation and has a much more narrow critical period than other V1 features.

These three features - ocular dominance, orientation, and direction of motion selectivity – rely on molecular guidance cues, endogenous activity, and natural experience in different ways. OD development begins through molecular guidances cues but requires spontaneous activity to mature normally; for the majority of cells, OR formation requires spontaneous activity for initial formation; and DR selectivity occurs only through visual experience. For this reason, and others, the work in this thesis involves features which are known to be affected by spontaneous activity - measurements and manipulations of orientation selectivity in chapter four and

binocular correlations in chapter five. Although the review presented here relies primarily on experimental measurements of topographic organization, as opposed to measurements of individual cells, a relationship is assumed where no other evidence is provided.

Although this dissertation stresses the effects and theoretical justifications of patterned spontaneous activity, there are clearly limitations to the role of this activity in development. Ocular dominance columns initially form through molecular guidance mechanisms, and subsequent activity was only needed for maintenance and plasticity during the critical period. Ringach's connectivity model (2004, 2007) demonstrates how V1 receptive fields and a functional topography could form based on the quasi-regularity of the ON/OFF center retinal ganglion cells in the retinal mosaic. Alternately, some aspects of V1 appear to not develop until after eye opening – e.g. a topography of direction selective cells. This review is meant to stress the way in which these stages – molecular guidance, innate learning, and visual experience – interact to form a mature V1. Certainly a number of molecular-guidance and experience-driven mechanisms are necessary for proper development, and even if rudimentary receptive fields and some topographies can theoretically form through those mechanisms alone, the simplicity, experimental activity manipulations, and functional benefits of endogenous activity suggest a substantial role in development. The visual system will eventually refine based natural experience, and work in this thesis will demonstrate that pre-experience receptive fields can refine using the same mechanism on simple patterns of activity.

Chapter 3: Modeling approaches to adult vision and visual development

The previous chapter reviewed the neurophysiology of the early visual system and the role of spontaneous activity. Here we summarize the computational approaches which help frame these experimental results. There are two main distinctions in the modeling approaches to be discussed; these are analogous to the two more abstract levels of analysis as discussed by Marr (1982). First, efficient coding strategies will be explained – specifically in how they relate to the early visual system. This approach involves understanding the overall objective for the neural code in V1 rather than the details of implementation. The abstraction of this level of analysis permits application in computer vision, and even application of the same coding algorithms to other sensory systems; the relation to coding in the auditory system will be briefly discussed. What is lacking in the literature is the application of efficient coding models to understanding early visual development and spontaneous activity – this is the impetus for the current work.

The second modeling approach is referred to here as neural modeling. It is computational in the sense that it is modeling by simulation, but is concerned, at some level, with how the specific algorithm is instantiated in the brain. These neural models attempt to reproduce the measured neurophysiological results for development, but without higher-level guiding principles like in efficient coding. As will be shown, in the case of development these neural models are underconstrained. Although some neural models of development include forms of spontaneous activity, the wide variety of these models does not allow any predictive value; the types of spontaneous activity vary as widely as the models which rely on them.

It should be stressed, as mentioned in the first chapter, that these two modeling approaches – efficient coding models and neural models (equivalent to Marr's computational and algorithmic levels, respectively) – are not competing models, but rather overlapping levels of analysis. The neural modeling approach can bridge between neurophysiological facts like local neural connectivity and higher-level properties like V1 receptive field maps. However, it can be difficult to construct neural models that also conform to a higher-level interpretation. For example, when we encounter difficulty characterizing neural responses in V2 and beyond, neural models may have too many degrees of freedom to explore these neural responses in a coherent, principled way. Similar to many other scientific explorations, we approach such unconstrained problems using abstraction. Before neural models were able to derive adult V1-like cells using natural inputs, research on sparse and independent coding demonstrated that V1 simple cell responses could be produced by an efficient coding of natural scenes. It appears that in V1 development, the historical trend from abstraction to implementation is reversed, with the consequence of a multitude of neural models without a high-level principled justification for model selection – principles like sparse or independent coding as detailed in the following section.

Sparse and independent coding in vision

The goal of early sensory processing is to efficiently encode our sensory experiences – this is a statement of the efficient coding hypothesis (Attneave 1954, Barlow 1969). In practice this computational approach involves maximizing a particular metric of efficiency. Such simple metrics can help explain properties of the adult visual system. For example, neural recordings of simple cells in early visual cortex show a preferred pattern of light and dark illumination. The V1 simple cell has elongated subregions of visual space where relatively bright or dark parts in the visual image will stimulate the

cell. Images of these receptive fields have been fit by functions like 2D gabors – planar sine waves within gaussian envelopes (Jones & Palmer 1987). These receptive fields can also be described as linear filters – a 2D map of scaling coefficients for image pixel luminance values which when summed together can predict neural response. These particular filter shapes describe a significant portion of the response of these neurons, but why these particular functions? Codes were optimized for a set of natural images with either sparse coding (Olshausen & Field 1996) or ICA metrics (Bell & Sejnowski 1997). The resulting codes produced filters that appeared qualitatively similar to the receptive fields measured in early visual cortex (e.g. 2D gabor functions). The similarity in codes measured from real neurons and codes derived from efficiently encoding natural scenes provides evidence that the visual code is adapted to efficiently code our natural, visual world. Neurophysiological recordings of adult V1 simple cells can give an approximation for what the neurons respond to (e.g. a gabor filter fit), but this efficient coding work is a normative description, demonstrating why those neurons respond the way they do.

In this dissertation, the terms sparse coding and independent coding will be substituted by the more generic term 'efficient coding', with an understanding that the particular form of efficiency referred to in this thesis is an objective more appropriate for neural responses in V1, as opposed to other efficiency metrics. A discussion of the particular merits of sparse coding versus ICA is orthogonal to their use here. To understand why only one coding objective was used in the research, it is important to note the practical as well as mathematical similarities in these two efficient coding approaches, which will be explained after a brief introduction to each technique.

Sparse coding is searching for a code where a given image is represented by a few

neurons firing strongly and a large number of neuron firing weakly or not at all. In other words, the histogram for a population of neural responses to an image should peak at zero and have a heavy tail; that is, sparse codes are those that produce a kurtotic firing rate distribution to a given class of images. The reasons for this goal include a minimization of metabolic costs, as spikes are metabolically expensive (Laughlin et al. 1998), and potentially improving learning by generally reducing the number of active neurons used to represent a given input. Sparse coding can also be understood by contrasting it with compact coding. If we are trying to minimize the size of the overall code used to represent an input (e.g. minimize the number of neurons in the brain), it is a compact code. A good example of this is compression on a computer using ZIP for files or JPG for images. The underlying code is 1's and 0's and the goal is to minimize the number of 1's and 0's in the new representation. Although it will not be discussed here, compact coding is the analogy that is often used for the type of coding which occurs in the retina to minimize the number of axons [See (Atick & Redlich 1992) and a critique by (Graham et al. 2006)]. In contrast to compact coding, human language is a prototypical example of the use of sparse coding. In language, the number of words is large, but we only use a very small number of them at any given time. The maintenance of a large vocabulary is less critical than shrinking the length of an utterance. V1 is thought to employ a similar strategy in part because the number of neurons in V1 is many times larger than the number of axons that make up the optic nerve – a ratio of over 200:1 in the macaque (O'Kusky & Colonnier 1982). Sparse coding may simply be a way to take advantage of such a large neural ensemble to represent visual information.

Related to sparse coding is independent coding, where the goal is to maximize the amount of statistical independence between neurons. One particular example is

independent components analysis (ICA) (Bell & Sejowski 1997). ICA has been used successfully in blind-source separation to, for example, separate sound sources in the cocktail party problem. By placing multiple microphones in a room, the mixture of sound sources can be separated based on their unique, independent statistical properties. Such a goal – finding the independent sources of input – would also be important in the early visual system. Quantifying statistical independence, especially with complex, high-dimensional data sets can be difficult, so the particular algorithm used can require a number of simplifications and assumptions; this can be seen in one form of ICA detailed in appendix A-1.

Sparse and independent coding are two distinct objectives, although in analyzing natural scene statistics they have been shown to produce qualitatively similar filters (Olshausen & Field 1996, Bell & Sejnowski 1997). Sparse and independent coding are not only related conceptually and in their results, but more rigorous mathematical derivations can show how they are related. According to the central limit theorem, the sum of a sufficiently large number of independent, random variables tends to a gaussian distribution. ICA implicitly uses this principle and searches for independent components by finding projections which are maximally nongaussian. In many ICA formulations, individual filters are sought that produce kurtotic firing rate distributions over a set of images. The goal of sparse coding can be formulated as trying to find a population of filters in which the firing rate histogram is highly kurtotic for a given image. These differences are equivalent to the terms lifetime sparseness and population sparseness, respectively (Willmore & Tolhurst 2001), or selectivity and sparseness, respectively (Lehky et al. 2004). There is not only a similarity in the maximization of kurtosis, but it is also known that for ergodic signals the two goals overlap. This is known to be the case for natural visual signals (Lehky et al. 2004). A

more rigorous mathematical treatment of the relation between sparse coding an independent coding is in appendix A-2. However, for the purposes of this work, it is only necessary to understand that the two approaches are related mathematically and produce similar neural filters. The exact ICA method used in the results section is outlined in appendix A-1.

It should be kept in mind that the efficient coding approach, for all its mathematical sophistication, embodies the idea that the ultimate goal of any neural representation is to be useful to the broader goals of the animal. The type of representation should increase the animal's evolutionary fitness. The more explicit and immediate objectives such as sparse coding or independent coding can appear somewhat removed from this ultimate goal without explanation. An objective like compression in the retina is beneficial given the limited size of the optic nerve. Such a strategy allows more data to be sent with fewer axons. The V1 coding objectives are efficient representations in part because they make the potentially relevant underlying information in the images more explicit. That is, neurons are representing the relevant sparse or independent sources of the data, and with such explicit representations, it can make learning new associations simpler. Other broader fitness reasons can exist for these objectives. The energy cost of neural spiking is relatively high, demonstrating the need for an efficient, sparse code (Lennie 2003). Calculations based on known measures of neural metabolism have demonstrated how a network of spiking neurons with a relatively small number of active neurons can use less energy than a network in which multiple neurons fire at intermediate rates (Attwell & Laughlin 2001). Explicit representation and metabolic efficiency are ecological justifications which give credibility to the approach beyond mathematical simplicity and ability to produce receptive fields.

The previous efficient coding techniques were aimed at understanding the adult visual system. That work primarily addressed why this code would form in an adult, and not how it develops in an individual organism. It could not address to what extent the visual code is hardwired by genetics versus adapted to the environment of the individual. Natural experience is necessary for proper V1 development, as has been studied by manipulations of experience in a series of visual critical periods (Bear 1995). The statistical structure of the environment can dramatically alter the visual code. This adaptation appears to maximize the efficiency of the code on the experienced inputs. Precisely how this adaptation occurs based on visual experience is not clear, however, a number of attempts at how the initial V1 code can form have been made, and will be discussed in the following section.

Neural models of V1 development

Although efficient coding models have not been of much application to visual development in the past, there is an extensive literature applying neural models to V1 development, as reviewed in this section. However, these models differ from adult efficient coding models not only in their level of analysis (Marr's "computational" vs. "algorithmic") but also on which of two features of V1 they focus. First, there are the receptive field profiles (e.g. the 2D gabor functions for simple cells) which are characteristic of V1, and have been produced by efficient coding models by adapting to natural scenes. However, in most efficient coding models, there is no prespecified arrangement of these neural filters. This is expected, since such an arrangement is not related to the computational goal, but rather the implementation of that goal. In contrast to this receptive-field focus, many neural models focus primarily on the development of the ordered, topographic arrangement of V1 neurons. The selectivity of neighboring neurons in primary visual cortex vary characteristically over

retinotopic location, orientation, ocular dominance, direction selectivity, and spatial frequency. These multiple dimensions have to be represented on the 2D surface of cortex; in order for neurons with similar features to be located near each other, unique overlapping organizations of these features occur. Many experimental studies have characterized and manipulated these maps. These neural models typically simulate neurons or groups of neurons which interact in physiologically plausible ways to produce cortical maps and receptive fields with expected physiological properties.

There are a number of similarities between the dozens of models available to form these maps – see Erwin and Shulten (1995), Miikkulainen et al. (2005), and Swindale (1996) for reviews. Here we discuss a few of the models to give an indication of the variation of approaches used. Linsker (1986) demonstrated orientation column (OR) and receptive field formation in a multi-layer model beginning with uncorrelated noise. Grabska-Barwinska and von der Malsburg (2008) demonstrated orientation column formation using recent experimental evidence of patchy, spatially periodic cortical spontaneous activity. The model of Miller et al. (1989) developed ocular dominance (OC) as well as orientation columns (OR). More current models have become even more ambitious in the combined development of feature maps. Bednar and Miikkulainen (2003) demonstrated direction selectivity (DR) to create a combined map (OR/DR). A later model combined these features (OC/OR/DR) using translated natural images (Bednar and Miikkulainen 2006). Carreira-Perpinan et. al. (2005) using the elastic net model included a spatial-frequency map (OC/OR/DR/SF), although the relation of their input to either natural stimulation or spontaneous activity is not clear. The use of general purpose algorithms like the elastic net or self organizing map algorithms relates to the general approach of trying to combine a high number of feature maps on a 2D cortical surface. In each of these models the goal was

to synthesize a cortical map and receptive fields which mimic known neurophysiology.

The first shared component of these models is that the neural connections are physiologically plausible: they rely on plausible synaptic modification rules – primarily Hebbian synapses; there are prespecified, and sometimes adaptive, short-range excitatory connections with slightly longer inhibitory connections between neuron in cortex; and there is a normalization of synaptic strength. Although early models had fixed, often circularly-symmetric lateral connections (e.g. von der Marlsburg 1973; Miller 1994), later models had adaptive connections form, which match the known, patchy connections between neurons of the same ocular dominance or orientation (e.g. Miikkulainen et al. 2005). Although there are unifying themes of physiologically plausible connections, there is a wide variety of mechanisms which can result from these basic building blocks.

The second component to most of these neural models is the presence of activity to drive refinement. These models show how different forms of spontaneous activity can produce a map representing primary visual cortex. However, the wide variety of neural modeling approaches also leads to a wide variation in the form of the spontaneous activity in these models. For example, von der Malsburg's initial model (1973, recently in Grabska-Barwinska & von der Malsburg 2008) used simplistic bar-like stimuli, Linsker's model (1986) begins with uncorrelated noise, Miller's model (1989, 1994) uses radially-symmetric functions that represent the amount of correlation between units, and Bednar and Miikulainen (2003) used 2D oriented gaussian blobs. Many such models rely on constraints in the neural connectivity (e.g. prespecified or adaptive lateral connectivity or dendritic field sizes) for receptive field

formation with less emphasis on the precise nature of the activity. As with many models, the necessary constraints to form V1-like receptive fields and cortical maps can either be in the network implementation or the particular statistics of the input, with many of these models scattered along this continuum. This lack of constraint leads to the production of a wide variety of network implementations, and a subsequent lack of understanding what necessary statistics of the image - spontaneous or natural - are needed to guide this development.

It should be noted, that there are also models which produce this topography not from spontaneous activity, but do so using natural images. In comparison to the previous models, these models tend to rely more heavily on the statistics of the input rather than precise connective properties. In these cases the models learn topography from natural scene statistics (Barrow et al. 1996, Hyvärinen and Hoyer 2001, Shouval et al. 1997, Weber 2001). For example, the topographic ICA model of Hyvärinen and Hoyer constrains the pooled responses of neighboring filters to be maximally independent, as opposed to maximizing the independence of the individual filters themselves. There is no direct efficient coding justification for such a constraint, but it does produce a cortical map which varies characteristically over orientation and spatial frequency, while relying primarily on the statistics of the input. For example, some of the previous forms of spontaneous activity mentioned in the preceding paragraph would not form reasonable filters using topographic ICA, as the statistics of the input are more critical in this case.

Between these two groups of topographic models, there is a subset of developmental models that have used natural images in addition to spontaneous activity; in the spirit of the innate learning principle, they apply the same learning method both before and

after visual experience (Burger and Lang 1999; Bednar 2002; Bednar and Miikkulainen 1998, 2004). Such neural modeling approaches provide the most direct relation to the work presented in this dissertation. Burger and Lang trained a hebbian binocular model with noise images followed by exposure to natural scenes; however, their postnatal learning stage completely rearranged the cortical map, which is known not to occur in animal development. The most extensive work on such combined modeling efforts has been done using the LISSOM framework. It has been applied extensively, with dozens of publications as well as a book primarily based on this framework (Miikkulainen et al. 2005). This neural model is at a large enough scale (approximately simulating cortical columns instead of neurons) to address inputs as complex as entire images. The model has also been extended beyond V1 to understand object segmentation, grouping, innate face preference, and other high-level perceptual phenomena. But most importantly, this map modeling approach has been applied to both intrinsic spontaneous activity and natural inputs – a critical component to employ in understanding an innate learning approach.

The primary goal of many of these neural models is to show how a 2D cortical map and receptive fields can develop in a physiologically-plausible way. This differs from efficient coding approaches in which the details of implementation are not specified. For example, topography is a necessary limitation for physiological plausibility where a neural implementation requires local connections, but topography is not necessary in an efficient coding model where only the high-level objective of the code needs to be specified. However, it is important to note not only the differences, but the intimate relation between the two modeling approaches. Although neural map models have simulated visual cortical development for decades, they have provided little constraint to inform us about the statistical nature of spontaneous activity. The efficient coding

approach applied to development in this dissertation relates the properties of spontaneous activity to receptive field formation, but leaves the critical component of V1 topography unaddressed. These do not need to be isolated approaches. Creating a neural map model which approximates an efficient coding objective would show the clear levels of analysis relationship between the two approaches. Similarly, efficient coding approaches could add spatial neighborhood constraints, as was done in topographic ICA (Hyvärinen and Hoyer 2001), to relate to the findings of neural map models, also blurring the distinction to demonstrate how they are related. Such bridging comparisons would focus efforts on the statistical properties of spontaneous activity and natural scenes, rather than particular and sometimes arbitrary constraints in the learning algorithm that are unique to the particular neural model chosen.

UNIT 2: RESEARCH RESULTS

Chapter 4: The initial, monocular innate learning model

The main theme of this chapter is a demonstration that a physiologically-relevant visual code can be learned through patterned, spontaneous neural activity. First, the spontaneous activity model created for this purpose will be related to currently known physiology as well as theoretical mathematical constructs (e.g. percolation networks). The filters learned by this activity will be compared to physiological receptive fields and filters derived from natural scenes. Later analysis will show how altering the parameters in this spontaneous activity model can lead to altered properties of these learned filters; this will show how slight manipulations, perhaps due to pharmacological manipulations or interspecies differences, can produce distinctly different receptive fields before eye opening. There will be a brief exploration of what the sufficient statistics are for receptive field formation, and the expected difficulties in a precise definition. Finally, additional results will demonstrate why such high-level, efficient coding statistical models are more beneficial in understanding our natural world than previous statistical approaches that relied primarily on simple correlation measures.

Methods: pattern generation, patch collection, and efficient coding approach

This section provides details of how the spontaneous activity patterns are generated, and how, precisely, images are coded to produce filters. Because the nature of this pattern generation is related to physiology, the details of pattern generation are particularly pertinent to the results and discussion. The specific patch collection and efficient coding procedures delineated here are the exact same used throughout this

dissertation regardless of the source of image patches - activity patterns, natural images, or non-physiological noise patterns. Although this dissertation does not favor a particular sparse/independent coding strategy, the fastICA algorithm (Hyvärinen 1999) was chosen for its simplicity, standardization in the literature, and computational speed. A brief explanation of the fastICA algorithm is in appendix A-1 and the particular parameters chosen are stated later in this section.

Pattern Generation: The pattern generation technique is a variation on site percolation (Aharony & Stauffer 1994). It is a simple, three parameter (p, r, t) model involving initiation and complete propagation of wave activity - thus the patterns used have no temporal component. On a square array of points, mark a random fraction 'p' of the points on the grid as potentially active. To initiate an active cluster, randomly select a location and activate all available points in a neighborhood radius 'r'. Neighboring potentially active points near the wave are activated if there are at least 't' active points within a distance 'r'. The wave is allowed to propagate until no more cells can become active. In the experiments here, waves were propagated on 256x256 pixel grids with circular boundary conditions. Although this fully explains how the patterns are generated using the three parameters (p,r,t) it does not address the interesting aspects of the behavior.

Percolation theory involves networks with t=1, often with r=1 as a typical example. When 'p' approaches a value known as the percolation threshold, 'p_c', the pattern of activity is known to be fractal – the image statistics appear similar at all scales. For example, when $p < p_c$, the activation terminates forming small clusters, when $p > p_c$ the activation spreads without bound leaving small holes without activity, but when $p = p_c$, both the activity and holes are nearly infinite in extent, leading to a fractal

interpretation. Images at increasingly larger scales become indistinguishable. For example, the following (p, r, t) triplets are known empirically, and in some cases analytically, to produce fractal images: (~0.592, 1, 1), (~0.407, √2, 1), (~0.288, 2, 1) (Malarz & Galam 2005). In our experiments for a given 'r' and 't' pair, p_c was approximated by finding the maximum derivative in the function of maximum cluster size to 'p' value. To obtain enough edge statistics on these waves, points were randomly chosen to begin wave propagation until more than 20% of the available points were activated – stopping when the last wave was allowed to fully propagate. Only the statistics of these final patterns were analyzed. All encoding was done by downsampling the image by 2 to minimize any local edge effects due to aliasing. Unless otherwise noted, 'r' was set to 3 for simplicity.

Patch collection and coding: Here we detail the specific coding procedure to find these filters. First, we generate and/or collect the images for sampling. Then we randomly select fixed-width patches from these images so the analysis is computationally tractable and we have enough samples. Next, the data is preprocessed ("whitened") to reduce the dimensionality, remove noise, and equalize the power spectrum of the image patches. This allows the ICA algorithm to work more efficiently, spending fewer computational cycles on the lower-frequency components. The exact mathematical description of this preprocessing step is explained in appendix A-3 in the discussion of filters versus bases. After this preprocessing, the fastICA algorithm, as explained in appendix A-1, is applied to find filters which maximize the statistical independence of the filters, producing a sparse/independent code.

For both natural images and activity patterns, image patches were randomly sampled –

excluding patches within a patch width of the border. Natural images were normalized to have pixel intensities with zero mean and unit variance over the whole image. Patches that are predominantly uniform in luminance/activity were excluded to avoid processing images with little statistical structure – specifically, patches with a pixel variance less than 0.16 were excluded. This is equivalent to a requirement that between 20% - 80% of the original downsampled sites were occupied. For each 256x256 image (downsampled to 128x128) up to 100 viable 16x16 pixel patches were selected. This was done until 10 000 patches were collected, which were then encoded using the fastICA algorithm (Hyvärinen 1999) using the 'tanh' contrast function. The dimensionality was reduced to 100 using PCA.

The activity model: parallels to physiology and percolation networks

The spontaneous activity patterns used in this chapter resemble the only 2D imaged spontaneous activity available – retinal waves; this is demonstrated in figure 4.1, with examples of our generated patterns in figure 4.1d. Beyond a visual resemblance, the pattern generation technique also abstracts from the general properties and parameters of current retinal wave models. It should be strongly noted, however, that this is strictly not a retinal wave model but is rather intended to be an abstraction of the critical features of relevant endogenous activity in general. Rather than a precise model, this method serves as an existence proof for how a simple method of pattern generation can be used in this stage of development; later in this chapter, figure 4.4 demonstrates this same premise using another method of pattern generation. This work concerns the statistical nature of the produced activity but also hypothesizes where it originates. In comparison to retinal waves, LGN/V1 spontaneous activity has a more direct influence on cortical receptive field formation, and is capable of producing patterns for binocular visual development. LGN and V1 activity have been

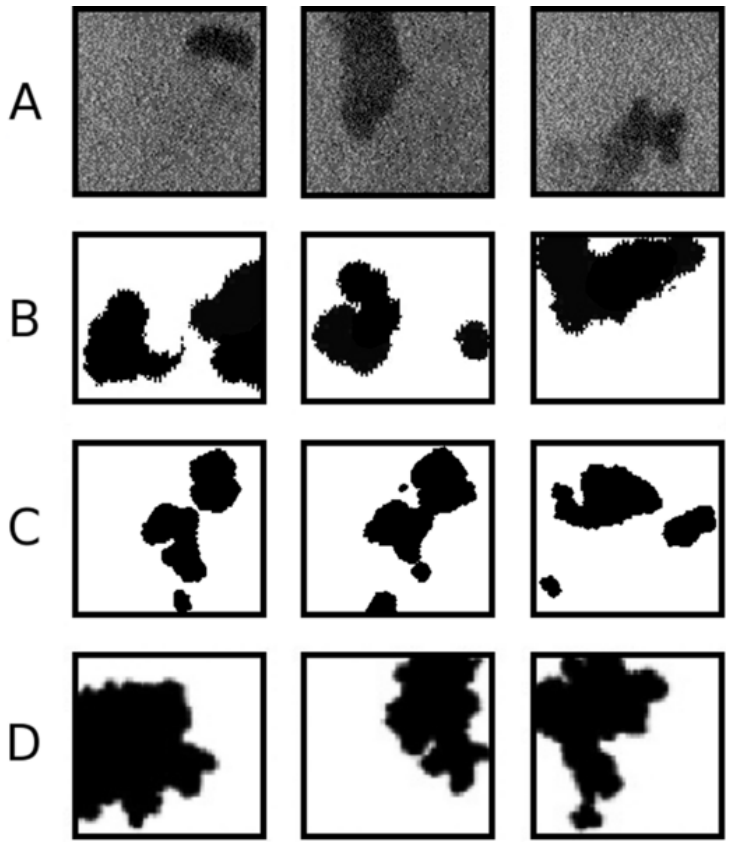

Figure 4.1: Experimental and theoretical 2D spontaneous activity images

(A) Experimental wave propagation: calcium imaging of a retinal wave (data as described in Feller et al. 1996). (B) Physiological model wave propagation: the ganglion cell layer activation of a retinal wave model (data from model described in Butts et al. 1999). (C) Physiological model wave extent: simulated retinal wave propagated to fullest extent (adapted from Godfrey and Swindale 2007). (D) Abstract model wave extent: a pattern generated by the technique used in this paper with parameters ($p = 0.55$, $r = 3$, $t = 6$) as detailed in the methods section.

experimentally characterized (e.g. Chiu & Weliky 2001, Weliky & Katz 1999), but are far less understood than retinal activity, thus prompting our analogies to retinal waves in this chapter.

The relevant statistical properties for an efficient innate learning strategy are present in a wide class of amorphous, wavefront patterns in which current models of spontaneous activity belong. The intent here is to demonstrate this generality, and avoid the pitfalls of selecting a particular physiological model, by using this abstract technique for pattern generation. This technique is a simple, three parameter model – a threshold, site percolation network model. Despite its abstract nature, this technique is analogous to known spontaneous activity patterns in generation and final pattern statistics. The retinal wave model of Butts et al. (1999) showed that wave propagation speed and termination were primarily determined by a 2D map of one summary variable, the local fraction of recruitable amacrine cells – similar to our variable 'p'. Their random variation of this parameter came from variations in cell refractory period, temporal dynamics from multiple waves, and influence of non-propagating spontaneous activity. Although the more recent Godfrey and Swindale model (2007) does not offer an equivalent summary variable, we believe a similar abstraction of local network excitability is equally possible. In the Butts model a neuron would only fire if a threshold of neighboring cells fired - similar to our 't' - while in the Godfrey model this threshold varied over time. Both models also had a fixed dendritic field size, analogous to our 'r'. Their parameters were chosen to match known physiological parameters such as wave size, speed, and frequency given neurophysiological constraints like known cell spacing and dendritic field sizes.

The pattern generation technique in this chapter was chosen for its theoretical simplicity and its abstraction of essential spontaneous activity features. Models common to the study of critical phenomena in physics, such as percolation models or the Ising model, have been used in artificial neural networks and understanding adult

retinal neuronal responses. Ising models, for example, have been adapted as artificial neural networks since Hopfield's network (Hopfield 1982). Recent work has also shown that Ising models are apt analogies for the maximal entropy and high-predictability of neural firing in the retina upon natural stimulation (Schneidman et al. 2006). Although the pattern generation technique we use is quite abstract, these other theoretical models have been shown to be relevant biologically by demonstrating desirable statistical properties. Additionally, the relation to more physiologically detailed models of spontaneous activity has been demonstrated, further warranting their application.

Learning receptive fields from the neural activity

After establishing a method of pattern generation, the next step in this research approach was to find if these patterns could be used to train an efficient coding system for natural vision. Sparse coding and ICA have been used to find approximately independent codes for natural images with resulting filters resembling those found in primary visual cortex, as shown in figure 4.2a-b. Figure 4.2b-d shows filters derived through efficient coding of activity patterns by using the fastICA algorithm (see appendix A-1). For these three panels, patches were taken from the various image sets using the image patch collection and coding as detailed in the previous methods section. Figure 4.2b shows filters derived from efficiently encoding natural images from the van Hateren image set (van Hateren & van der Schaaf 1998). To test if endogenous patterns of activity have similar statistical properties to natural scenes, one might ask if an efficient coding of activity from more physiologically precise models is capable of producing similar V1-like filters. To show this, retinal wave images from figure 2 of Godfrey and Swindale (2007) (examples of such images are shown in figure 4.1c) were efficiently encoded. The resulting V1-like filters from this

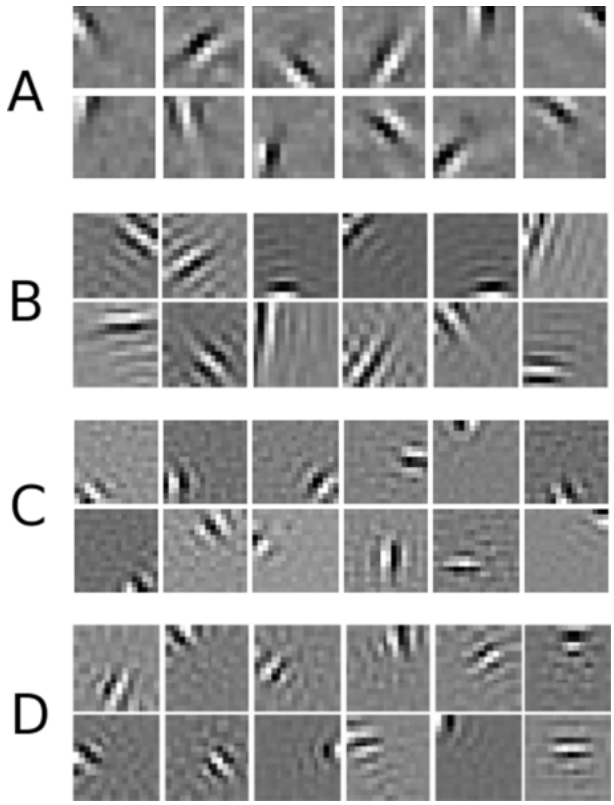

Figure 4.2: V1 simple cell receptive fields derived through an efficient coding of natural scenes and spontaneous activity patterns.

(a) Receptive fields from sparse coding: basis functions derived from natural images (algorithm as described in Olshausen and Field 1996). (b) Receptive fields from ICA: filters derived from natural images (algorithm as described in Hyvarinen 1999). For panels (b,c,d) the same patch collection and efficient coding algorithm was used as detailed in the methods section of this chapter and appendix A-1. (c) Receptive field filters derived from images of simulated retinal waves as in figure 2 of Godfrey & Swindale 2007 - a few examples are in figure 4.1c of this dissertation. Patch size for this data corresponds to approximately 0.3mm. Refer to the text for the implication of this result. (d) Receptive field filters derived from our generated patterns with parameters (p=0.7, r=3, t=8)

data are shown in figure 4.2c. Although an efficient coding of activity patterns produced by this model does qualitatively produces physiological filters, it is important to demonstrate that these images are embedded in a larger class of

amorphous, wavefront-containing patterns capable of producing relevant filter properties. The question of whether or not the activity comes from a particular model - or even from the retina vs. the LGN/V1 - is important, but in this chapter we stress the necessary statistical properties independent of the particular source. To demonstrate this, we generated a set of images from our abstract pattern generation technique with the resulting filters shown in figure 4.2d for comparison. Note how the general statistical structure found in natural scenes, our abstract patterns, and more physiological models of spontaneous activity all produce filters resembling those found in V1.

There are noticeable differences between receptive fields in both the initial and final stage of refinement by spontaneous activity. First, it is known in the ferret that initial molecular guidances cues appear to start the receptive field refinements with minor orientation selectivity (Chapman & Stryker 1993) and ocular dominance (Crowley & Katz 2000), as opposed to the random initial starting state of the derived filters. Although this clear difference in initial state may influence the resulting receptive fields, we believe the main difference of a nonrandom start state for the properties considered here is faster development toward a shared optimum. A more important difference between physiological and derived filters are the visible differences between receptive field profiles. For example, the produced filters have more lobes than physiological filters, which typically have between two and four – this would subsequently produce a much smaller spatial frequency bandwidth than physiological filters, however, we believe this deviation is only particular to the specific efficient coding method used – other encoding methods may provide similar minor variations in the produced filters. Also, as noted in the next section, the activity parameters can greatly vary the derived filter properties. Also, minor inconsistencies in filter

properties should be expected, as the precise objective function for V1 is not clearly known (e.g. sparse vs. independent coding) and we are using a limited, linear basis set – an inaccurate modeling simplification, even for linear-like V1 simple cell receptive fields.

Properties of the visual code depend on the precise activity parameters

Variations in the generated visual filters based on changes in the parameters of the activity model were investigated in one illustrative case. Although not strictly a property of physiological spontaneous activity patterns, the focus was on scale-invariant patterns in this model so sampling was done at $p=p_c$ (along the phase transition boundary) as shown in figure 4.3a. Approximate scale invariance is a property shared with natural images (Field 1987). In this case, it also allows neurons with limited dendritic fields to produce large-scale patterns with statistical properties that are relatively consistent across a wide range of scales. We also chose this sampling as a mathematical convenience so results would not depend upon a defined scale of analysis. Note that known spontaneous patterns, such as retinal waves, are clearly not on this self-similar boundary, but may be considered close; some species (e.g. ferret – Feller et al. 1996) have limited wave sizes while others (e.g. chick – Wong et al. 1998) cover large areas of retina often terminating at the edges.

To further demonstrate the ability of these amorphous, wavefront patterns to generate physiological filters, sets of images along the phase-transition boundary were generated. Filters derived from these samples are shown in figure 4.3c. A qualitative difference between the gabor filters is visible, and one aspect of these filters was analyzed. Orientation bandwidth was chosen because it is a well-defined, physiologically measured parameter. Seven parameter gabor filters were fit to the

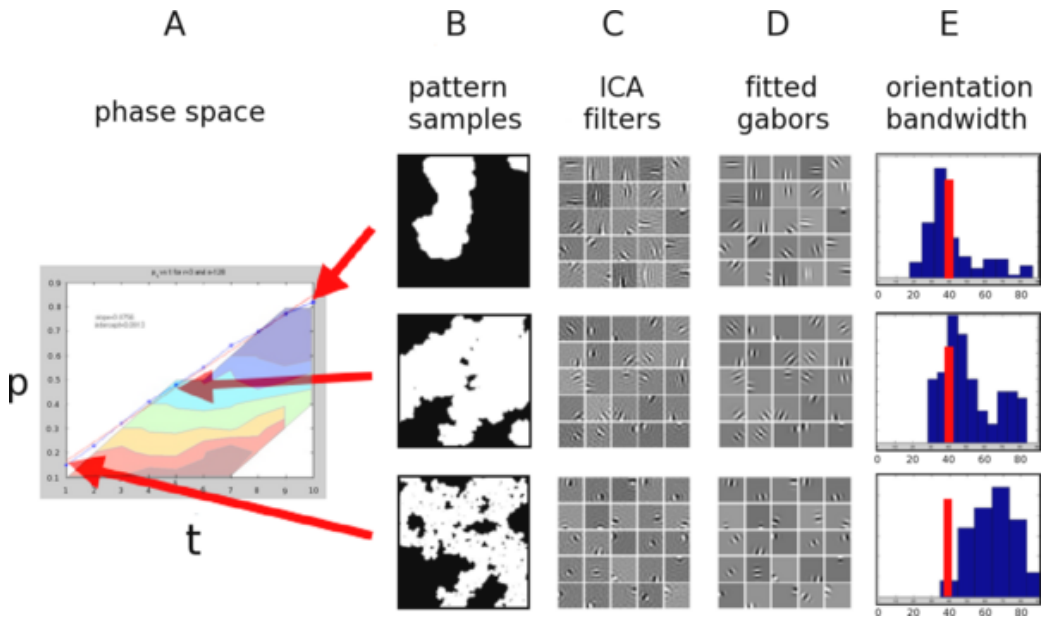

Figure 4.3: Summary of pattern sampling and analysis demonstrating relevant variation in derived gabor filters.

(a) Phase plane with 'r' fixed at 3. The curved line in the plot indicates the phase transition boundary as detailed in the text. The transparent color contours below the phase transition line indicate the trend for the median orientation bandwidth in that area of the plane. (b) Sampled patterns from (p,r,t) space near the critical percolation threshold - (0.15,3,1), (0.48,3,5) and (0.83,3,10). (c) The 16 x 16 pixel derived ICA filters. (d) Seven parameter gabor fits of those filters. (e) Histograms of the gabor orientation bandwidths in blue compared to the physiological median in red.

16x16 pixel derived filters. After this fit, the parameters of the gabor fits were used to find the orientation bandwidth, with histograms of these fits shown in figure 4.3e, along with the primate physiological median (De Valois et al. 1982). The area below the phase transition boundary for this parameter was also coarsely explored; the transparent color contour in figure 4.3a indicates how the median orientation bandwidth changes in this region. Note for $p < p_c$ a manipulation of 'p' is more effective at changing the orientation bandwidth than 't' – one indication of how models such as this one could lead to testable predictions of pharmacological

manipulations. However, we do not intend to make direct comparisons to physiological filters; such comparisons are difficult. Receptive field properties can depend on the precise encoding objective (e.g. sparse coding vs. ICA) and even the particular parameters of the learning model. The receptive field properties also change substantially after initial experience – e.g. orientation and spatial frequency bandwidth decrease as newborn macaques age (Chino et al. 1997). Even with this simple generation technique and imposed self-similarity constraint, a significant variation of filters can be produced. This variation demonstrates one way a spontaneous activity method can adjust local parameters to affect global pattern statistics and more closely resemble properties of adult physiological filters and natural scene efficient coding filters.

What statistical properties in activity promote V1-like filter formation?

The previous section demonstrated that it is possible to vary the properties of the derived gabor filters by varying the parameters in the spontaneous activity model; however, it is fundamentally the statistical properties of the spontaneous activity patterns, independent of how they were generated, which guide V1 refinement in this paradigm. This leads one to ask, what precisely are these statistical properties? Unfortunately, there is no simple set of concepts or measurements which can easily delineate this boundary. For example, in the last section the images were sampled along the percolation threshold, where the edges varied from smooth to jagged. This can be quantified by a variation in fractal dimension, but there are many other ways to generate edges of a particular fractal dimension, each of which alters a number of other image properties that can affect the derived sparse/independent code. Although there is no direct measurement of image statistics that can fully encompass the statistical structure on which sparse coding algorithms rely, it can still be helpful to

explore the space of potential patterns, as some measurements can provide an intuitive understanding of the necessary image structure.

Figure 4.4 explores the boundaries of a subset of images capable of producing V1-like filters upon efficient coding. Unlike the previously encoded images, these were produced through a well-defined, but distinctly non-biologically plausible method - thresholded noise images. Initial noise images were produced with a randomized phase spectrum and an amplitude spectrum using the chosen slope in a log-log plot of amplitude versus frequency. Both the slope of the amplitude spectrum as well as the number of thresholded levels were varied. The jagged vs. smooth edge-like structure of two-level images can be quantified through the fractal dimension. The fractal dimension is a quantity which varies between 1 (e.g. a straight line in a 2D plane) to 2 (e.g. a completely space-filling convoluted line in 2D) providing here a rough indicator for the smoothness of an edge. For the noise images, the fractal dimension of the edges was measured by the box-count method (as referenced in Schroeder 1992). The three images shown in Figure 4.4a were generated from noise with negative exponents in the amplitude spectrum of 1, 1.6, 2.2, and 4. The corresponding fractal dimensions are 1.80, 1.49, 1.27, and 1.19, respectively (with the maximum confidence interval in the estimate at 0.05) providing a relatively wide range to explore the impact of fractal dimension on the resulting filters.

The results of encoding with ICA using the same methodology as in the previous figures produced filters which are shown in figure 4.4b. Qualitatively, images with an amplitude spectrum slope near -2.0 and two threshold levels (white/black) most resemble the previously generated activity patterns and similarly produce the most physiological filters. As the edges lose structure, either by becoming more

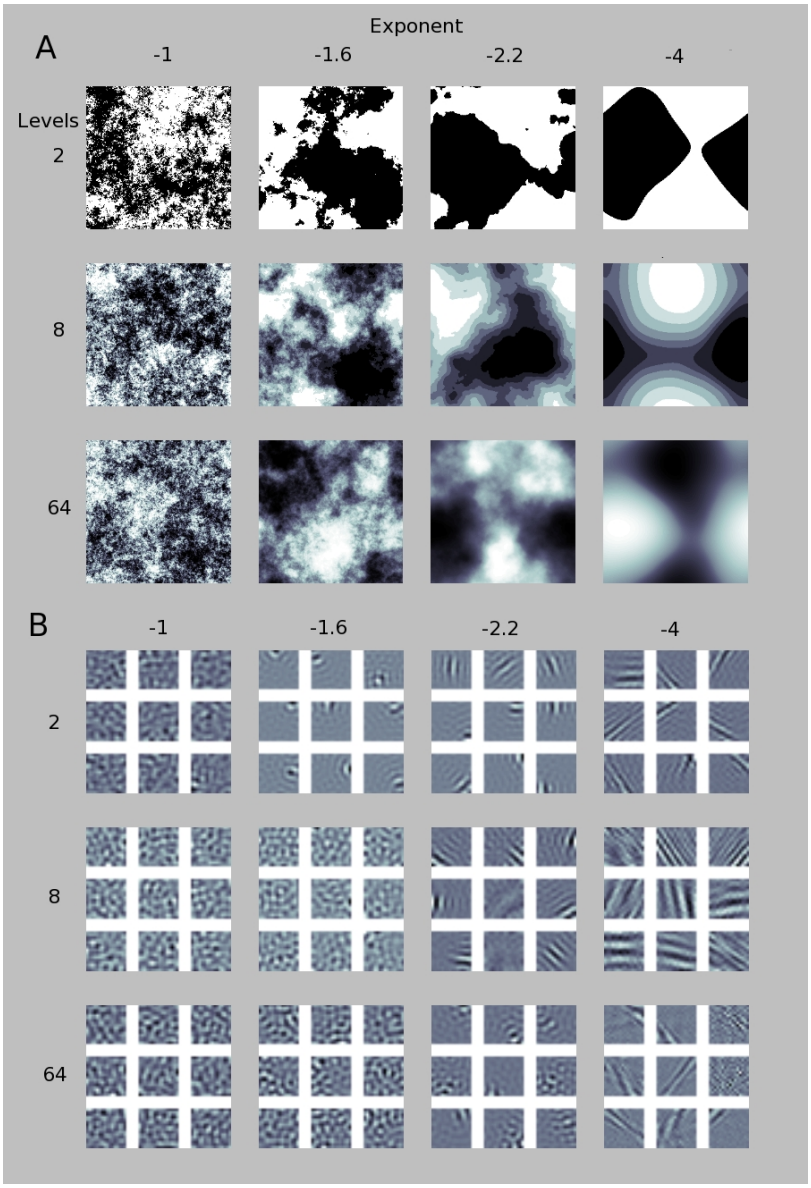

Figure 4.4: Efficient coding of a related pattern set: thresholded noise

Patterns were generated using thresholded noise to explore the statistical properties upon which sparse/ICA coding would produce V1 receptive fields. A) 256x256 pixel patterns produced from thresholded noise where the power spectrum slope is varied horizontally in this figure, while the number of threshold levels is varied vertically. B) 16x16 pixel filters produced by efficiently coding those images. Note, the parameters which resemble both physiological and modeled spontaneous activity produce filters which best resemble physiological filters. Additional explanation in the text.

jagged/convoluted on the left of the figure, or more blurred toward the bottom, the algorithm becomes less capable of discerning structure. When the edges are very straight, significant structure is captured but only in the form of highly elongated, unlocalized receptive fields. Figure 4.4 demonstrates that V1-like receptive fields are not only possible from biologically-plausible generated patterns, but any patterns with similar statistical structure.

The importance of higher order statistics

Current models proposed to explain pre-experience cortical receptive field development rely primarily on hebbian learning mechanisms and pairwise correlations in the input (second-order statistics). These approaches do not address the higher-level statistical structure for receptive field formation related to efficient coding. Although sophisticated hebbian-based models are technically capable of achieving arbitrary levels of complexity (and can even implement sparse coding strategies in complicated configurations as in Falconbridge et al. 2006) we note that most hebbian models rely solely on these pairwise correlations. In figure 4.5, a fundamental difference is addressed between second order and higher order correlations with respect to relevant statistical structure and receptive fields. The term 'pairwise correlations' refers to the fact that such statistics are based on the simple linear correlations of pairs of pixels for luminance or neural activity. Equivalently, the term 'second-order' statistics deals not with the analysis of histograms of luminance values (first-order statistics), but with the analysis of pairs of pixels. The point of figure 4.5 is that techniques which rely on these simple statistics are not capturing/concerned with a great deal of potentially relevant statistical structure. Note that uncorrelated noise (“white” noise) has no second order or higher order statistical structure, so techniques that rely on pairwise correlations (e.g. PCA) or higher-order statistics (e.g. ICA) do not produce filters with

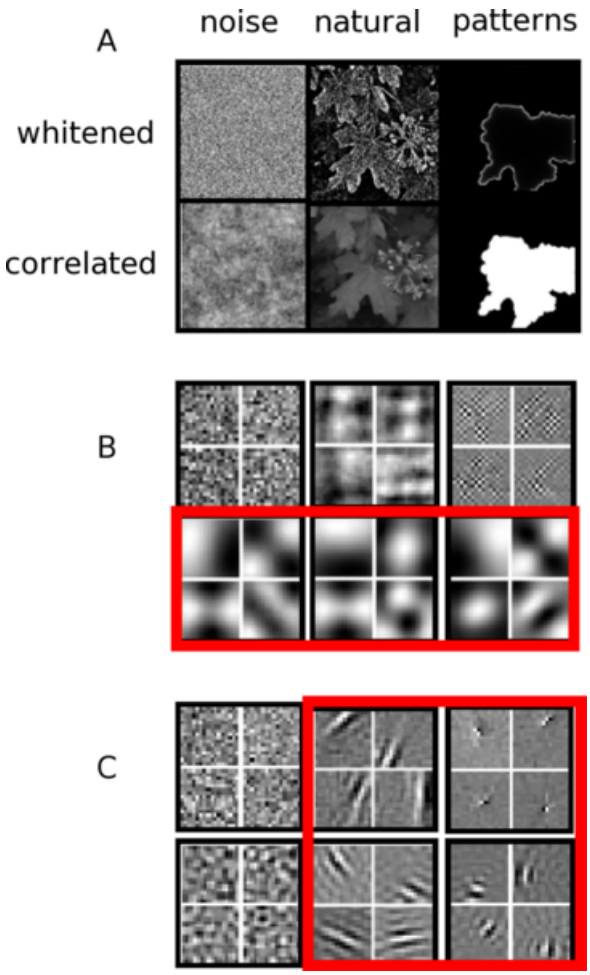

Figure 4.5: Information relevant to filter formation goes beyond simple, pairwise correlations

(a) Examples of training data for efficient coding. Whitened images were obtained by flattening an assumed 1/f slope in the Fourier amplitude spectrum. (b) PCA bases from each of the six data sets in (a). Note, algorithms which rely on pairwise correlations alone (also known as second order image statistics) only find structure for receptive field formation in correlated data [red rectangle] although much of the useful structure still visible in the decorrelated ("whitened") images is not captured. Also note the receptive fields in this case are not localized. (c) Representative filters from the same image sets using ICA. Note that the filters from the whitened and unwhitened, natural and wavefront-patterned data qualitatively resemble receptive fields [red square], whereas unstructured 1/f noise does not produce equivalent filters - unlike the results of pairwise correlation-based measures. (note: differences in whitened ICA filter sizes are primarily a product of the 1/f assumption)

discernible structure. In patterns with only the same second order correlations of natural scenes - as in the random 1/f patterns ("pink" noise) - PCA can produce filters, but these filters are not localized within the patch. Also, because the higher-order statistical structure is not present if these 1/f noise patterns, ICA does not produce V1-like filters. For natural and our patterned images in this figure, we have removed much of the the second-order correlations by a procedure to flatten the approximately 1/f slope in the Fourier amplitude spectrum. This removes the correlations in images that have an approximately 1/f slope, as was shown to be the case for natural images (Field 1987). When we whiten the images by removing the pairwise correlations, PCA bases resembling receptive fields are by definition unable to form, and we see that natural images as well as the wavefront patterns still retain important image structure. This whitened structure, reminiscent of line drawings, is efficiently encoded using ICA. Also note that for these image sets the ICA filters are inherently localized within the filter patches. However, encoding using PCA will not produce sufficiently localized filters without the use of additional constraints. Whichever encoding scheme is used, it should be noted that the generated wavefront patterns have both pairwise correlation structure as well as sparse, edge-like structure used by ICA. If only correlations were necessary to prepare the visual system, there are a number of even simpler ways to create these correlations without the additional wavefront, edge-like structure. This additional, higher-order structure can be exploited by the visual system to guide receptive field formation and maintenance. The fact that it may exist in both spontaneous activity patterns and natural scenes suggests that a shared method of receptive field adaptation may rely on these higher-order statistical properties.

Chapter 5: The binocular model and early depth perception

The last chapter laid the groundwork for extensions of the innate learning paradigm. The first model was spatial and only monocular. In this chapter, we explore the extension of this model to include an additional ocular layer. Although the change to the activity model is relatively minor, this opens up new analysis techniques and comparisons to physiology. First, binocularity is the most obvious reason why the necessary activity for innate learning cannot be situated in the retina. The links will be made between the activity model and what is known about the physiology of LGN/V1 spontaneous activity. Second, the resulting patterns will also be efficiently encoded to develop binocular filters. Image patches and filters derived from the activity patterns will be analyzed, specifically noting the disparity distributions and their relation to young primate binocular disparity distributions. Third, a framework will be proposed for evaluating the model through its ability to support early stereopsis. A physiological model that combine depth information from multiple disparity-sensitive filters will allow for simple depth-perception judgments given a stereo pair of images. This permits an evaluation of innate learning on an issue of direct perceptual and evolutionary relevance.

Model motivation and relation to LGN/V1 spontaneous activity

The reason for modeling multiple ocular layers is presented in this brief section – see chapter two for a more thorough treatment of the relevant physiology. The majority of experimentation on imaged, correlated endogenous activity in the visual system has focused on retinal waves. There are some important reasons why retinal activity cannot be the pattern used for innate learning of a V1 code. One immediate objection is that the waves are not binocular. Days after birth, primates have an adult-like

proportion of disparity selective cells (Chino et al. 1997) and retinal waves are not correlated between the eyes to promote such selectivity. Another criticism is that in some animals, like the ferret, retinal wave activity only minimally overlaps in time with the onset of V1 activity (Wong 1999). Also, the retinal waves of the ON and OFF-center cells are more independent at later stages (Wong & Oakley 1996) but the responses of these cell classes are intimately related in V1 neurons. Despite the relative popularity of studying retinal spontaneous activity, such concerns have turned many researchers away from considering the possibility of an innate learning role for spontaneous activity.

Here we draw attention to a form of spontaneous activity that does not suffer from the many deficiencies of retinal waves as a training pattern – activity originating primarily in the lateral geniculate nucleus (LGN). Although much less is known about this activity, it may provide new insights into the early development of neural response properties in the visual cortex. Endogenous activity in the LGN clearly overlaps in time with V1 responsiveness (Wong 1999). The activity is also correlated between eye-specific layers, as well as ON and OFF-center cell layers, as shown by the work of Weliky and Katz (1999); their figures showing the electrode activity over time provide evidence that the activity appears as a wavefront moving across these layers. By all accounts LGN endogenous activity appears better suited as a potential source of innate training patterns.

There are potential difficulties in providing a clear assessment of LGN spontaneous activity. First, it is more difficult to experimentally characterize LGN waves through electrode penetrations than the analogous characterization of retinal waves using calcium imaging. Also, proper LGN activity requires V1 feedback; there is no

correlated spontaneous activity between eye-specific layers in the LGN without feedback (Weliky and Katz 1999). The alleged source of training patterns for V1 may more correctly be called dynamic LGN/V1 activity, making a label for the source of such activity more complicated than a simplistic, one-area label. This of course, does not preclude the functional separation between an innate training pattern (LGN/V1 waves) and efficient coding technique (e.g., sparse coding/ICA), which by their nature need not be in physically separate locations; all that is required are separate functional roles. Note that the normative, efficient coding models here do not exclude the possibility of neural models with activity and adaptation occurring within V1 (e.g. Grabska-Barwinska & von der Malsburg 2008), although in this work we consider the combination of the LGN and V1.

Methods: extension to the binocular activity model

The model used is a variant of the initial threshold percolation model, but with an additional ocular layer and an additional parameter for the probability of activity transmission across these eye-specific layers. See chapter 4 for a description of the initial three parameter (p,r,t) monocular model; in that model, neighboring cells transmitted their activity with 100% fidelity. In the binocular model this certainty of transmission occurs within ocular layers; however, activity reaches corresponding neighbors across ocular layers with a fixed transmission probability – hereafter labeled 'trans'. This strategy reflects the fact that communication between ocular layers in the LGN is limited by required feedback from V1 and Weliky & Katz (1999) measured weaker correlations between ocular layers than by neurons within the same layer. The additional binocular parameter permits a range of low to high correlations between eye layers.

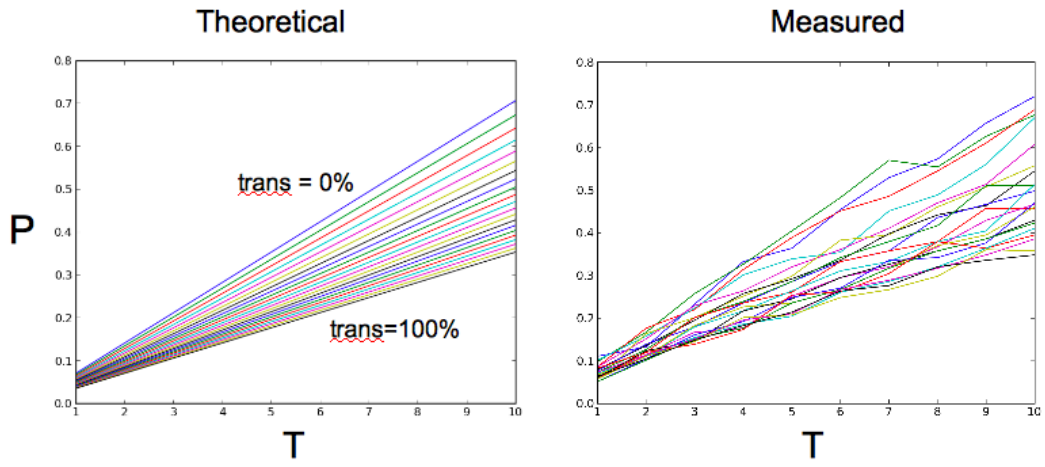

Figure 5.1: Theoretical and experimental percolation thresholds with r=3.

The two plots above show how the percolation threshold varies in the binocular model with a change in model parameters. Generally, as 't' increases, 'p_c' increases. Also, as the transmission probability between ocular layers increases from 0% to 100%, the necessary 'p_c' is lower. The rationale and equation used for the theoretical construct is in the text.

In many of the results the value of 'p' had to be adjusted to remain near the percolation threshold when the transmission probability between ocular layers was changed. As can be seen in figure 5.1, the measured percolation threshold varies nearly linearly with a change in 't', and the slope of this relationship changes with the interocular transmission probability. Decreasing the threshold number of neighbors needed to fire or increasing the fidelity of transmission between neighbors across ocular layers leads to an increase in the percolation threshold. This trend is modeled analytically by equation 5.1 below, which calculates p_c from the other three binocular model parameters.

(eq. 5.1) $$p_c = t / [(\pi r^2 / 2) (1 + trans)]$$

This analytical approximation for p_c is compared to measured values in figure 5.1. This formula is derived from the point at which one half of the neighborhood for a potentially activity cell would have an average of 't' potentially active neighbors. Without meeting this threshold a passing wave, which would often envelope only half

a dendritic neighborhood, would not be capable of significant propagation. Note that this formula is only a valid approximation for large neighborhood radii; at small radii there is an alias effect as the number of neighboring cells can change significantly with only a small change in neighborhood radius.

The following two figures, discussed in the next section, are the result of analyzing patches taken from running the four parameter binocular model to completion. Step-by-step instructions for activity pattern propagation and termination are detailed in the methods section of chapter 4, with the addition of the additional ocular layer and transmission between layers as mentioned here for the binocular model. Binocular filters produced by efficient coding use the exact same algorithm and parameters as chapter 4, with only the expected doubling of the initial patch dimensionality to accommodate the second ocular layer. Patches to be analyzed were taken from corresponding portions in the two different eye layers. For figure 5.2, 't' and 'r' were set to 1 for simplicity. The transmission probabilities ('trans') were 0%, 5%, 20%, and 80% with corresponding 'p' values of 0.55, 0.50, 0.45, and 0.35 respectively. The patterns for figure 4.3 used a neighborhood radius 'r' of 3 and a threshold number of neighbors 't' of 2. The low, medium, and high correlation conditions had inter-ocular transmission probabilities of 0%, 10%, and 50% respectively with corresponding 'p' values of 0.22, 0.17, and 0.13, respectively. The particular parameter values in both figures were chosen primarily for a variation between high and low correlation between ocular layers with no motivation for a particular choice of 'r' or 't'; 'p' was chosen to be near the percolation threshold.

Analysis of the binocular activity model

Figure 5.2 demonstrates an important result of this model – an adult-like disparity

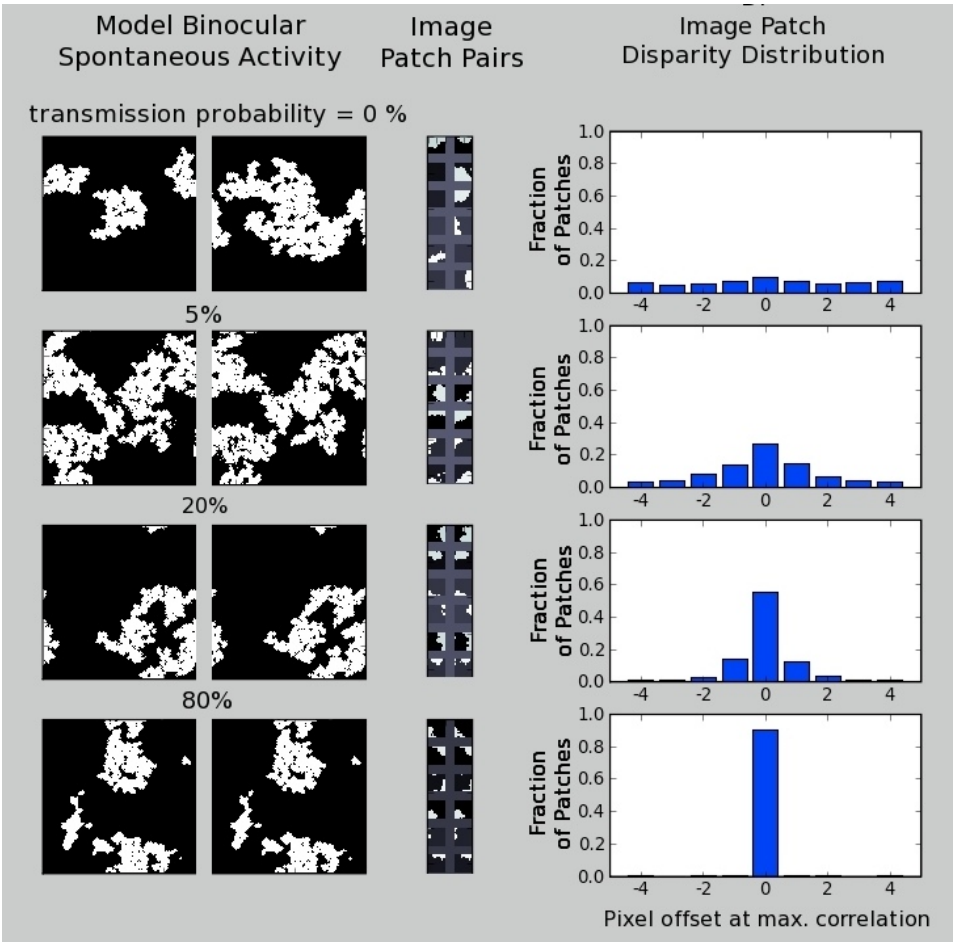

Figure 5.2: Binocular correlations and image disparity in LGN spontaneous activity.

Modeled binocular spontaneous activity is presented showing varying amounts of correlation between the eye-layers. The amount of correlation demonstrates the potential range in image disparity distributions. On the left are example images of activity in the two different eye-layers. The transmission probability represents the probability that activation in the model crosses to the other eye-layer; higher values cause higher correlations between eye-layers. In middle, example image patches extracted from corresponding locations in the two different eye layers. Note, large-scale correlations in the images may not be as evident in the smaller-scale collected image patch pairs. The plots on the right show image patch disparity distributions. Each graph represents the horizontal offset in pixels where the two image patches are maximally correlated.

distribution at weak correlations between ocular layers. Corresponding image patches across eye-layers were cross-correlated horizontally to find the horizontal offset that produced the maximum correlation between the left and right patches. A histogram of these maximum offsets represents a disparity distribution for the activity images. As can be seen, at zero correlations between the eye-layers the offsets are almost uniformly distributed. This is a case where binocular neurons would clearly not form. In the case of extremely high correlations binocular neurons form, but essentially the same image is present in both ocular layers; the disparity distribution becomes too narrow for depth discrimination using binocular disparity. Between these extremes, partial correlation between eye layers promotes a more physiological disparity range. Recall, these partial correlations between eye layers have been shown in LGN spontaneous activity and are produced by V1 feedback to the LGN (Weliky & Katz 1999); these correlations may relate to perinatal adult-like ocular dominance and binocular disparity distributions (Chino et al. 1997, Maruko et al. 2008). In this way, LGN appears capable of spontaneous activity that can help refine binocular receptive fields in V1 before eye opening.

The various correlations between ocular layers have a predictable effect on the resulting filters from an efficient coding of these images. Figure 5.3 shows that when there are no correlations the filters are monocular. With high correlations the filters are equivalent between the eyes and unable to provide disparity information. Only with an intermediate amount of correlation are filters produced which are localized but of differing location and phase. Although figure 5.2 only analyzes the image patches, as opposed to the produced filters, various image disparity distributions are produced; the intermediate distributions are indicative of adult-like disparity distributions found in postnatal primate V1 as early as two-weeks after birth (Maruko et al. 2008),

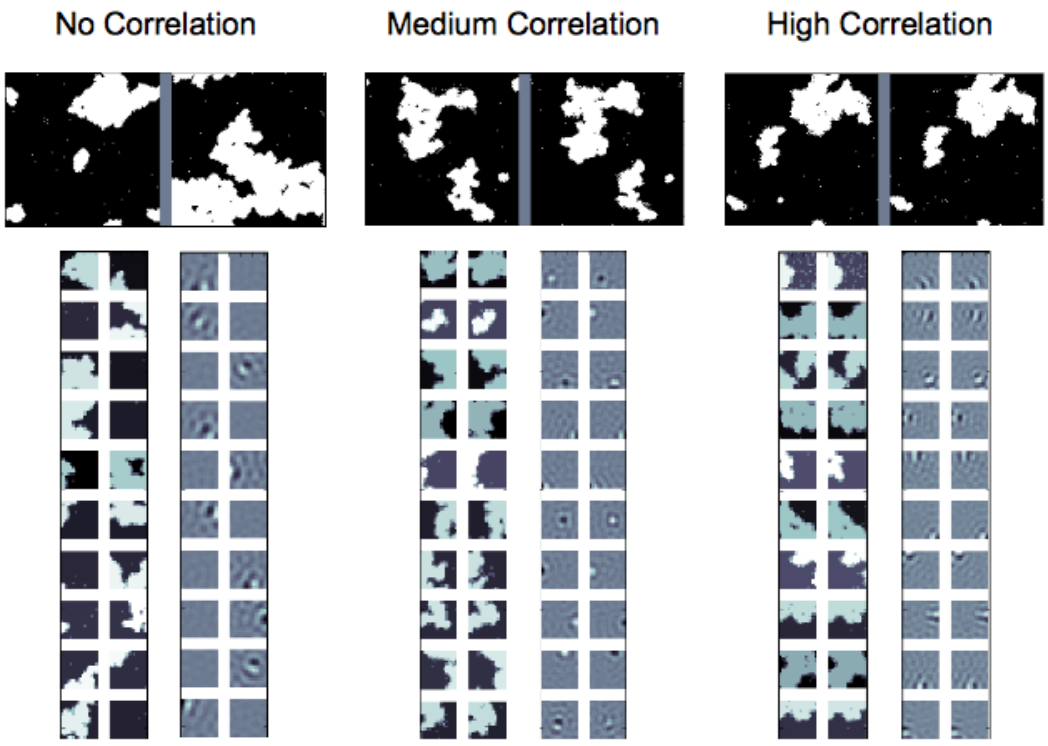

Figure 5.3: Filter formation at various eye-layer correlations.

For each of the three sets of correlated activity patterns there are example image and derived filter sets. Top: an example of the overall activity in the two eye layers. Left: example image patch pairs taken from corresponding locations. Right: filters resulting from an efficient coding of the image patches. Particular pattern generation parameters are found in the methods section of this chapter.

however earlier work by Chino et al. (1997) also showed an adult-like distribution of ocular dominance near birth. Examining figure 5.3 we see that filters produced by intermediate correlations between eye-layers may have a distribution of disparities, but appear to be equally binocular, with little or no eye-preference as evident in V1 ocular dominance columns. This is not a unique theoretical result; ICA applied to natural stereo images also produced filters with a strong tendency for equal binocularity rather than biased preference for one eye (Hoyer & Hyvarinen 2000). A possible reason for this discrepancy in ocular dominance distribution, as mentioned in a V1

developmental model by Berns et al. (1993), is that a period of monocular input, prior to or interleaved with binocular input, is necessary to establish a physiological ocular dominance distribution. We know that early in ferret development there is an overlap between retinal wave activity and LGN and V1 spontaneous activity (Wong 1999). During this time, the monocular patterns from retinal waves may promote monocular receptive fields in V1. Also, the extent of monocular cells due to molecular guidances cues prior to spontaneous activity is not known. The amount of transmission between ocular layers could also be varied to promote a particular balance of monocular and binocular cells. Although not explicitly modeled, any combination of these methods are expected to be capable of manipulating the ocular dominance distribution in the innate learning paradigm to match the adult-like ocular dominance distribution found near birth.

Overall, the binocular activity model demonstrates that through the addition of one parameter to the initial activity model, the effects of varying the correlation between eye layers can be explored. This has been demonstrated in the image patch disparity distributions, the effects on filter formation, and the theoretical results showing how this affects the percolation threshold. In the next section, we further explore the implications of a binocular model to an aspect of perception – depth discrimination.

Depth discrimination with the binocular activity model

Instead of evaluating the binocular model using binocular disparity distributions or receptive field profiles, the effect of innate learning can be demonstrated through application to depth perception. We have seen that a physiological distribution of disparity sensitive cells may be derived using spontaneous activity. The stimulus-driven activity of a set of receptive fields derived from this model can be combined to

reconstruct the disparity from a stereo pair of images. A given set of binocular filters can be tested by analyzing their ability to find the original depth map from a stereogram. In the innate learning paradigm, spontaneous activity parameters can be chosen to maximize the system's ability to make depth discriminations. As a demonstration of the idea, an optimal set of filters was applied to the autostereogram (e.g. "magic eye") in figure 5.4 and produced a depth map from the local image disparities that matches the map used to generate the stereogram. The same evaluation technique can be applied to optimize the parameters of binocular spontaneous activity. Clearly animals which are able to use stereopsis earlier in life would be at an advantage, especially for visually precocial young. This method can such a salient advantage in natural selection to the form of binocular spontaneous activity.

Methods: the depth discrimination algorithm

The autostereogram of figure 5.4 was produced using the algorithm from Tyler and Clarke (1990). A 1/f noise pattern was used instead of a random dot pattern for the tiled background to help match corresponding points at multiple spatial scales for easier viewing. The autostereogram was chosen for presentation purposes as it is easy for people to binocularly fuse the images compared to a traditional stereogram image pair. For initial processing, the autostereogram was separated into a left and right-viewed image – slightly offset views as would be seen by a subject when the two eyes are at the correct vergence angle. For a given binocular cell, the stereogram images were filtered by the corresponding left and right linear filters. The distribution of disparity judgments for a particular cell was found by horizontally cross-correlating the left and right bases corresponding to those filters. These individual binocular cell disparity judgments were summed and weighted by the product of the left and right filter activity for that cell for a particular location on the image (the reasoning for the

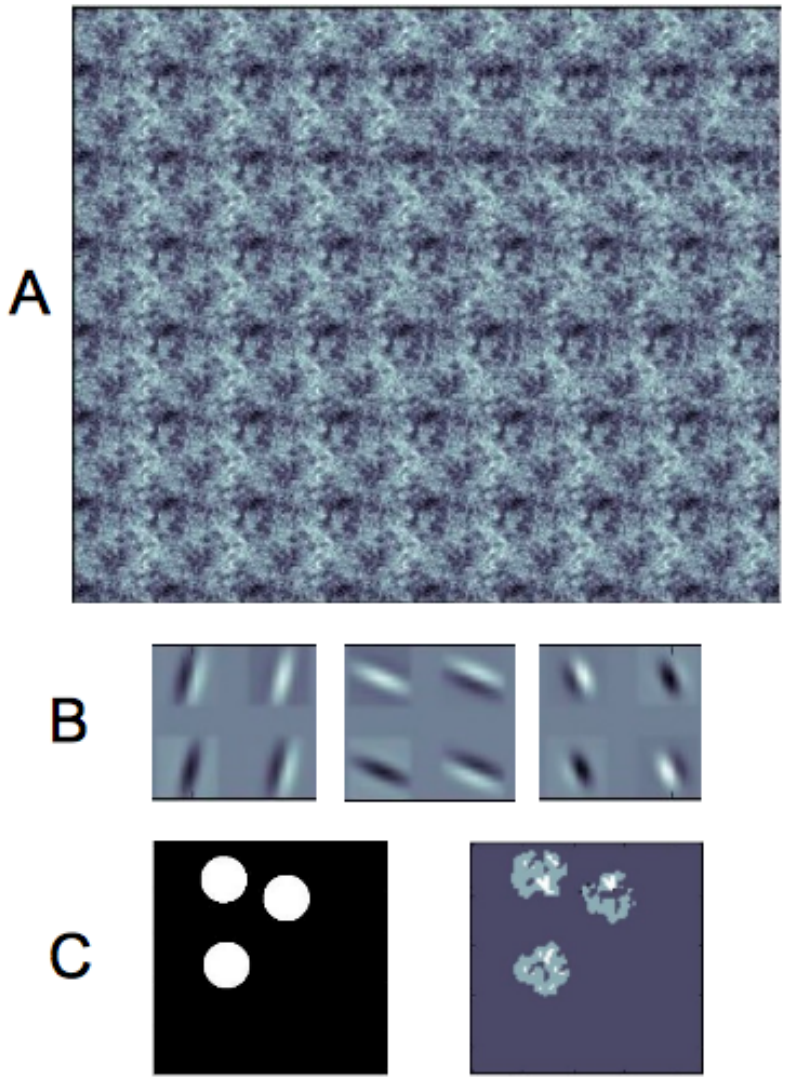

Figure 5.4: Depth discrimination using binocular receptive fields.

A demonstration of the disparity judgments possible with receptive fields of various binocular disparities. A) an autostereogram with the depth map given in C. B) an example of 3 pairs of binocular receptive fields used to analyze the stereogram – the filters here are quadrature pairs with randomly distributed phase offsets between the ocular layers. C) the left image is the original depth map used to generate the stereogram. The right image is the depth map derived from applying the filters to the stereogram.

product is discussed in the next paragraphs). The overall disparity judgment for a particular location in the image was taken as the disparity with the maximum value based on the weighted sum judgement over all the cells. Although Qian and colleagues describe a multiscale algorithm to sharpen the image only one scale was used here, as it is only for demonstration purposes.

Abridged derivation of the algorithm: Due to the offset in filter position and/or phase between the left and right receptive fields, simple and complex cells demonstrate disparity tuning (Ohzawa & Freeman 1986a,b). Binocular simple cells are sensitive to disparity, but the responses are also highly modulated by the phase and precise location of the stimulus in the same way as monocular simple cells are sensitive to the precise phase of a sine wave grating. Complex cells are not as sensitive to the precise location and approximate a disparity detector well for such an early stage of processing. Qian and colleagues (Qian 1997, Qian & Zhu 1997) created a model capable of combining the information from disparity-selective complex cells to form surface disparity judgments. In that work, they used theoretical complex cells made up of a sum-of-squares quadrature pair of two equally phase-shifted binocular simple cells – a binocular equivalent of the complex cell energy model (Adelson & Bergen 1985, Pollen & Ronner 1983). Simple derivations from these theoretical cells showed that the disparity selectivity comes from the products of the left and right filter responses for each individual cell; ihis result is demonstrated in equation 5.2 below. The energy model complex cell response is represented on the left of the equation, and expanded on the right. As L_1/L_2 and R_1/R_2 are quadrature pairs, the first two pairs of terms are insensitive to phase/exact position. The final product term provides the necessary disparity information for a theoretical complex cell.

(eq. 5.2) $(L_1+R_1)^2 + (L_2 + R_2)^2 = (L_1^2 + L_2^2) + (R_1^2 + R_2^2) + 2\,(L_1\,R_1 + L_2\,R_2)$

This approximation is valid for a quadrature pair of simple cells that are part of a theoretical complex cell derivation, but the linear filters that form through spontaneous activity are not in quadrature pairs. Qian also demonstrated analytically that equivalent disparity tuning is possible with a set of binocular simple cells that are not quadrature pairs; the only requirement is a uniform distribution of phases (as opposed to the 90 degree offset of quadrature pairs) for binocular cells of the same phase-disparity. In practice, however, quadrature pairing and uniform distribution are not possible. Here the tacit assumption is made that the phase/position offset between binocular cells is distributed randomly enough to negate the stimulus phase and precise position dependence of individual cells. With a limited set of filters which are not precisely gabor-like, perfect disparity discrimination is not the goal. In fact, there are many other algorithms capable of better stereo depth judgment that are not performed in very neurally plausible ways. The goal here was simply to use a physiologically plausible stereo algorithm to demonstrate an additional evaluation method for derived binocular filters.

Conclusions of the binocular model extensions

An abstract spontaneous activity and efficient coding model allow for straightforward extensibility of the concept of innate learning to more complex problems. The extension of the initial monocular model to binocular layers was relatively simple - one additional parameter and the formation of the filters by efficient coding used the same coding algorithm as the initial model. Although the technical changes in the activity model and efficient coding model between this chapter and the previous chapter were minor, the impact of the results and analysis were significantly more extensive. The model in this chapter introduced binocular aspects of early spontaneous activity that clearly cannot occur in the retina – thus showing the need to

study LGN/V1 activity. Also, the partial correlations found between ocular LGN layers in physiology were shown to lead to an adult-like binocular disparity distribution for the efficient coding filters. The resulting filters match certain properties in the physiology of binocular cells near birth, but may also be useful for early stereopsis. By using an algorithm derived from theoretical models based on complex cell physiology, disparity judgments from the binocular linear filters can be combined over the entire stereogram. Figure 5.4 showed that the algorithm can produce a reasonable representation of depth. This depth discrimination framework is one potential way of demonstrating an immediate advantage of innate learning. By refining based on patterns of activity before the eyes open, binocular cells can be improved, and the capability for depth discrimination early is life is possible.

UNIT 3: DISCUSSION AND IMPLICATIONS

Chapter 6: Implications of the current work

There are two direct benefits from this normative approach to development – a better understanding of known neurophysiology and an additional methodological approach to research of visual development. Although previous neural models have used forms of spontaneous activity, the particular nature of the activity varied considerably and provided little insight into the amorphous, wave-front properties found in both retinal and LGN/V1 activity. The approach presented here relies on higher-order statistical structure shared between natural scenes and spontaneous neural activity; this constraint allows us to generalize beyond natural scene statistics in a way that explains observed qualitative properties of spontaneous activity. This constraint required the use of a spontaneous activity model which resembled retinal waves qualitatively and reproduced certain properties of LGN waves like partial binocular correlations. The main finding is that the currently measured experimental properties of spontaneous activity may be a direct result of an innate learning paradigm.

Although the models in the previous two chapters explain current neurophysiological properties of spontaneous activity, they also provide a roadmap for additional properties that should be measured. For an example discussed in detail later, the 'p' vs. 't' phase space map in figure 4.3 may provide certain insights into pharmacological manipulations and their effect on receptive field characteristics. Most importantly, this work addresses new insights into the unique statistical properties of spontaneous activity. Many of the neural models previously used relied on pairwise correlations, and those correlations do not explain the additional structure of wavefront-containing

amorphous patterns (as was mentioned in figure 4.5). Higher-order statistical measures, such as kurtosis as used in sparse or independent coding, may be an additional driving force in learning by both endogenous activity and external experience. Finally, the most compelling advantage of this modeling strategy is parsimony. The two distinct stages of development, before versus after eye opening, may share the same learning algorithm. This insight can be approached from an efficient coding level of analysis, as done in this dissertation, or at the level of neural models of V1 map formation - mirroring the 'computational' vs. 'algorithmic' levels of analysis of Marr (1982). The insight of innate learning provides informative links between disparate research areas, which will simplify the overall task of understanding the early visual system.

Retinal and LGN/V1 neurophysiology: experimental predictions

This dissertation heavily relies on known aspects of retinal spontaneous activity, however, the source of spontaneous activity for innate learning in V1 (the LGN with V1 feedback) may only marginally involve the retina. The reason for this discrepancy is the amount of knowledge about retinal vs. LGN and V1 spontaneous activity. Retinal waves are easily accessible and imaged, they are relatively isolated from the rest of the nervous system, and the detailed structure of the retina is known. This has led to the production of computational models to simulate these wave properties (Butts et al. 1999, Godfrey and Swindale 2007). The same level of detail is not available for LGN and V1 activity; the LGN is not easily accessible or imaged. It is much less isolated than the retina not only because of feedforward connections from the retina and to V1, but also because of V1 feedback. Few studies have performed recordings of LGN spontaneous activity during development (Weliky & Katz 1999). Given this lack of characterization, there are no models of LGN spontaneous activity as there are

for retinal waves (Butts et al. 1999, Godfrey & Swindale 2007). Because of this, the models of Unit II are primarily based on known retinal wave physiology. However, it must be stressed that these are not retinal wave models, a fact that is clearly demonstrated by the extension to binocular activity.

The innate learning principle can be understood without referring to particular models, but modeling approaches force an answer to questions about the nature of spontaneous neural activity. The binocular model provides an example of a particular model choice requiring experimental verification. The disparity distribution in receptive fields learned from innately generated activity is largely a function of the distribution of alignments between wavefronts in the two eye-layers. The extent to which this misalignment is along the horizontal axis between the eyes versus distributed in both vertical and horizontal directions is unknown. In our modeling effort we did not specify any bias for the activity, which allowed for an equal amount of vertical as well as horizontal displacement. This displacement can be measured experimentally by analyzing spontaneous activity recordings. These measurements also have implications for receptive field properties at birth. If the innate learning paradigm holds, activity with horizontal and vertical displacement between eye layers will produce receptive fields with similarly unbiased displacements. This unbiased displacement is not the case for adults; after significant natural experience the displacements are primarily horizontal, as would be dictated by the statistics of natural binocular experience. Cumming (2002) found that the the variation in preferred horizontal disparities was 3.6 times larger in the horizontal direction compared to the vertical direction in primate V1. The presence of a variation in preferred vertical disparities may be either from natural vertical disparity (for example, objects closer to one eye are vertically larger in one eye than the other) or from an unbiased initial point

in development. This can be confirmed by measuring either the spontaneous activity patterns themselves or analyzing the newborn primate disparity bias. LGN activity has not been sufficiently analyzed to answer this question and the newborn primate bias has not been tested – receptive field displacements for individual cells early in life are difficult (but not impossible) to measure due to low spatial frequency monocular properties of cells at that time (Chino et al. 1997). As evident in this example, such a modeling approach forces questions, even with the simple binocular activity model used, helping to guide possible future experiments.

The modeling effort here is not only useful for making predictions about normal physiology, but also providing insights into potential experimental manipulations. For example, the exploration of the 'p' vs. 't' phase space map can be related to small pharmacological manipulations, as both parameters have physiological equivalents. There is a debate between the necessity or sufficiency of spontaneous activity – to what extent is it permissive vs. instructive for a V1 visual code. Most experimental manipulations dramatically alter this activity by abolishing or greatly overstimulating the neural activity (Stryker & Strickland 1984, Lowel & Singer 1992, Weliky & Katz 1997, Chapman et al. 1986, Chapman & Godecke 2000). Pharmacological manipulations may be able to explore this instructive role using less extreme manipulation. It was shown in figure 4.3 that below the percolation threshold orientational selectivity varies with a change in 'p'. The change in 't', the threshold number of active neighbors needed to activate a neuron, had less of an effect but permitted a wider range of 'p' values by increasing 'p_c'. The prediction is that changes in the threshold mechanism for spreading a wave are not as important as variables which affect the fraction of recruitable cells during wave propagation, at least with respect to the measured variable in that figure – orientation bandwidth. There are two

pharmacological routes to produce these manipulations in the retina, where the mechanism is well understood. If the probability of amacrine cells firing independently (which initiates wave propagation) could be decreased, this would have the effect of increasing the 'p' value, as more time would occur between waves for more amacrine cells to become recruitable. This would have the immediate effect of increasing wave speed and size, but would also likely produce waves with lower fractal dimension, as is the case in our example patterns in chapter 4. Alternately, a variation in the threshold parameter, 't', may be possible by depressing the ability of amacrine cells to transmit impulses between cells. In both cases, it is possible to relate even the highly abstract models used in this thesis to testable predictions through manipulations of the pattern generation parameters.

An additional prediction of the model in its current form is that spontaneous activity may not be ideally suited for receptive field formation at multiple scales. In appendix A-3, figure A.2 shows that filters produced by an ICA coding of natural images are present at multiple scales/spatial frequencies, however, patterns resembling those from the spontaneous activity model fail to produce filters at multiple spatial scales. One experimental prediction is that these waves may not produce filters at multiple scales. Although a mixed noise image model shows that some variation in filter width can be achieved through mixing multiple types of waves, the results do not demonstrate a clear ability to produce filters at multiple scales, thus prompting further work.

We know from physiology that at least in primates, although a number of binocular properties are near adult levels (Chino et al. 1997, Maruko et al. 2008) the initial monocular properties are low in spatial frequency. The development of the spatial frequency map has been indirectly studied in the developing cat by measuring the

presence of cytochrome oxidase blobs (Murphy et al. 2001). It was found that after 2 weeks postnatally blobs begin to form, and this initial formation is independent of visual experience. It should be noted that color-responsive cells are also more abundant in V1 blobs making this result less clear as a differentiation of spatial frequency. Initial research in V1 development involved orientation selectivity and ocular dominance, however, these simulation results point toward an interesting limitation of spontaneous activity in reference to spatial frequency. In the same way that direction selectivity appears, at least in the ferret, to require natural visual experience (Li et al. 2006) spatial frequency organization, at least as the current simulation shows, may also require visual experience.

As can be seen, these models allow for testing at a much finer scale than simply abolishing or overstimulating. Results from retinal wave research can be incorporated into LGN/V1 spontaneous activity models to make testable predictions. The predictions we have discussed include the possibility of horizontal disparity bias, sensitivity to specific pharmacological manipulations on activity parameters, and the possibility that spontaneous activity may produce only one spatial scale of filter. Over time these models may not only be further validated, but more importantly can help guide new neurophysiological measurements.

Efficient coding: the importance of higher-order spatial structure

The unique aspect of this innate learning approach is the use of efficient coding principles to learn from spontaneous activity. This has led to a focus on receptive field properties rather than V1 cell topography in this dissertation. Previous neural modeling approaches to development relied on a mixture of decisions on neural connectivity and spontaneous activity modeling (see reviews by Erwin & Shulten

1995, Swindale 1996). These neural models have focussed on the properties of lateral connections and intracortical adaptability rather than the instructive role of spontaneous activity. With so many free parameters, it was possible to create receptive fields with impoverished activity models – e.g. uncorrelated noise (Linsker 1986), lines (von der Malsburg 1973), or gaussian blobs (Miikkulainen et al. 2005). Because of the many ways that the cortical network parameters could affect the results of neural activity (specifics of lateral connectivity, degree of initial specification, balance of inhibition/excitation...), it is more difficult to disambiguate the effect of neural network specification from the form of spontaneous activity in these models. In effect, if the constraints could be built into the network rather than the statistics of the spontaneous activity then there was no need to seek a parsimonious 'innate learning' solution.

As the models became more developed and constrained by our knowledge of physiology, there were research approaches which attempted to link development before and after initial experience. Many models did this implicitly – for example, Miller's model (1989, 1994) relied on correlations within and between eye-layers, which can be expected of spontaneous activity as well as natural signals. However, two models explicitly applied the same neural learning model to both images of spontaneous activity as well as natural scenes (Burger and Lang 1999; Bednar 2002; Bednar and Miikkulainen 1998, 2004). Although these models used an innate learning strategy – using the same learning method for endogenous as well as external activity – these neural models provided enough constraints on receptive field formation that impoverished activity models (e.g. those with only second-order correlations) could also produce acceptable receptive fields.

The general methodology in both physiological literature and these neural modeling approaches has been to consider the pairwise correlational statistics in the input patterns. The work on LGN (e.g. Weliky & Katz 1999) and V1 spontaneous activity (e.g. Chiu & Weliky 2001, 2002) has primarily measured the pairwise correlations between various electrode positions. Also, the Miller model (1989) relied on simple circularly-symmetric correlation functions between positions. Miikkulainen et al.'s LISSOM model (2005, for a review), as well as many other neural map models, require only correlated input. The pairwise correlations are not the only important statistical structure. For example, as shown in figure 4.5, these simple correlations can be removed by 'flattening' the power spectrum of an individual image. Images with the pairwise correlations removed exhibit easily recognizable structure similar to line drawings. This is evidence that these correlations are not the most important statistical structure to capture. Additionally, the reliance of developmental models on pairwise correlation statistics may indicate that these models are ignoring important higher-order statistical structure sought in current efficient coding models.

Sparse coding attempts to find codes in which only a few neurons fire strongly while the rest are nearly silent. In other words, the search is for a code which produces a kurtotic distribution in firing rates to a given image. Mathematically, kurtosis is a fourth-order cumulant as evident in the formula for kurtosis found in appendix A-1 (mean, variance, and skew are first, second, and third-order cumulants respectively). As mentioned in chapter three and the appendix, ICA also seeks a kurtotic distribution in firing rates. In fact, as a first step in these algorithms, the pairwise correlations are often removed in a process known as “whitening”. Figure 4.5 demonstrates this issue fairly clearly. We can identify whitened and unwhitened images, and the efficient coding algorithms learn an efficient code for both classes of images equally well.

However, algorithms which rely on pairwise statistics, like PCA, not only fail to pick up the statistical structure in whitened natural images, but also produce receptive fields in situations which intuitively have little structure – like 1/f noise images.

Most importantly, however, the reliance of previous models on pairwise correlations is not only out of sync with current work on efficient visual coding, but may also mislead efforts to find the relevant statistical properties for innate learning. For example, even in figure 4.5 we can see that the PCA receptive fields are not localized. The reason why many neural models can create localized receptive fields is not based on the statistical properties of the input, but rather by hard-coded limitations of local neural connectivity. When the incoming and lateral connections of V1 neurons are limited to local connectivity, localized receptive fields necessarily occur regardless of the statistical properties in the input. Applying a localization constraint opens the door for modeling approaches which may disregard the statistical properties which allow efficient coding models to form localized receptive fields. Higher-order statistics are likely to play an important role in development, as they are critical to understanding adult visual function from a computational perspective. They have been understandably ignored in the neurophysiological literature due to the distant level of abstraction, but should not be so overlooked in computational approaches. This will be revisited in the section on neural models in the final chapter.

Simplicity: the conceptual and evolutionary advantage of parsimony

In addition to developing the proper models for neurophysiology and statistical interpretation, there is an additional advantage to this innate learning approach – parsimony. Although understanding the approach requires a background in physiology as well as computational modeling, the insights and links promote a

simple, more holistic understanding than both fields can provide separately. First, development before and after eye-opening appear inextricably related, as the same learning algorithm may apply in both stages. Second, neural models of V1 receptive fields and map formation can be related to efficient coding techniques of understanding adult vision – bridging work between two disparate computational areas of visual system research. Finally, and most importantly, this parsimony is not only helpful for conceptual purposes, but can be of direct benefit to the animal as well. The simpler developmental strategy may provide an animal with evolutionary advantages – simpler genetic coding needed for setting up the visual system, faster development by exploiting the time before the eyes open, and possibly speeding up the rate which mutations may advantageously affect the species.

There is a tendency to divide research areas into distinct divisions so that researchers can understand and communicate effectively within a particular domain. Normally this leads to increased effectiveness when connections within domains are high and between domains are low. However, even in nearly ideal circumstances, there will inevitably be links which are lost due to these divisions that can provide critical insights. There is a traditional focus on what is innate versus what is learned – nature vs. nurture. Although it is well known that there are many aspects of animal development which depend on both, there is one boundary in this debate which remains relatively clear. After eye-opening there may be a mixture of learning and genetic, innate influences but the time before eye-opening is primarily innate. Of course, exceptions are made for light coming through the eye-lid (Akerman et al. 2002) and malnourishment from the mother, but in general eye-opening appears to be the point where significant effects due to the animal's environment begin.

In this work, however, we show that the same learning algorithm can be applied on natural images as well as innately generated patterns of activity to refine a visual code. This simplifies our understanding prior to eye-opening in a number of ways. First, the particular patterns of spontaneous activity are not as random as previously thought – the particular statistical properties may relate to the statistics of natural scenes insofar as they can equivalently refine V1 receptive fields through learning. Second, it is the same learning algorithm. What we know of learning in critical periods during experience may apply before natural visual experience. Also, as will be mentioned in the final chapter, the use of the same learning algorithm before and after eye-opening may not be the only link between these stages - V1 spontaneous activity also plays a significant role in adult learning and development. There are a number of reasons why these two stages of development appear to be one stage with a change in inputs.

Another bridge that this work creates to simplify our understanding of development is between two different computational approaches – here labeled efficient coding and neural modeling. The neural modeling approach has been dominant in development, focusing on creating V1 maps and at the same time V1 receptive fields. Efficient coding has been used to understand adult vision, justify it from an ecological perspective, and apply the approach to computer vision algorithms. This work helps to bring these two approaches together, by showing that a popular efficient coding algorithm can be applied to better understand development.

The parsimony of the approach is much more than a benefit to our understanding; such simplicity brings evolutionary advantages to animals that employ such a strategy. First, a simple method of development requires simpler genetic specification. When that same visual learning system for natural vision can be used on innate patterns, the

only additional genetic encoding necessary for innate learning is for pattern generation. As we have seen in chapters four and five, this can be done quite simply. Rough molecular guidances cues are necessary for initial connectivity, but refinement using spontaneous activity is a fairly economical way of preparing the visual system. For completeness, it should be noted that the simplicity of the learning strategy is not an argument supporting an exclusive role at any stage in development. It is possible that any one of a number of overlapping development strategies ranging from molecular guidance cues to altricial, supervised natural experience can also play prominent roles. It has been claimed that instead of being understood as an elegantly simple design, the brain may employ a number of these distinct, but functionally overlapping adaptive strategies – as Marvin Minsky aptly stated "the human brain is a kludge." There is some benefit to even part of the "kludge" being simply specified and implemented, perhaps even to the benefit of freeing genetic resources for other competing and overlapping developmental strategies.

The other advantage of the innate learning approach is the speed of learning. Certain animals, such as deer, must be visually precocious and able to follow their mothers soon after birth. By having the V1-equivalent of natural visual experience before eye-opening, the visual system can be mature enough to accomplish this task. Spontaneous activity brings the neural code closer to adult performance, thus decreasing the amount of time the early visual system must adapt to the visual environment. This faster development is also critical as later stages of visual development depend upon a relatively stable lower-level visual system.

The less intuitively obvious, but also potentially advantageous effect of a parsimonious developmental strategy may be its effect on the dynamics of natural

selection. Not only can the result of an evolutionary process be advantageous, but also the speed of the evolutionary approach itself. In order to learn similar neural connectivity for V1 at eye-opening, evolutionary adaptation of molecular guidance cues alone would likely be significantly slower. A simple, innate learning approach has the benefit of producing a significant change in network connectivity at eye opening with fewer evolutionary steps. This understanding is in line with similar self-organizing approaches for complex systems. Valsalam, Bednar, and Miikkulainen (2007) showed that even for handwriting recognition systems, convergence in an evolutionary algorithm is faster in a network using a self-organizing system for an initial bias. It was demonstrated that self-organizing innate learning systems allow faster natural selection to improve a neural network for applied vision. The analogy is that improving the speed of advantageous evolutionary mutation provides a competitive advantage for species that use this strategy of development.

Clearly, the modeling of innate learning using knowledge of retinal and LGN/V1 spontaneous activity provides new insights into development. Such models not only allow us to understand development under normal function, but also upon manipulation – like the pharmacological manipulation in the 'p' vs. 't' phase plane and its relation to orientation selectivity shown in figure 4.3. This work can not only guide better neurophysiological understanding, but also demonstrates the need to focus on the correct statistical approaches. The higher-order statistics on which efficient coding techniques rely are often not captured by models which only need pairwise correlated inputs. Although this work touches on a number of separate areas, it is in the pursuit of simplifying our overall understanding. What was previously two stages of development appears to be one stage with only a change in inputs. This work also begins to bring together different levels of analysis by showing that efficient coding

principles can also inform developmental models in ways that may benefit neural models. However, this innate learning strategy ultimately only exists because of its evolutionary benefit to the animal. Such strategies not only allow simpler genetic encoding and faster development for an individual animal, but also easier, faster incremental change by mutation for the species overall. For these reasons, the innate learning paradigm helps us not only make sense of various stages of development and levels of analysis, but also demonstrates the direct benefit for the animal from this method of development.

Chapter 7: Potential directions for future research

There are three broad directions for future work on the principle of innate learning. First, a spatiotemporal model can supplement the monocular and binocular spatial models. Second, the evolutionary implications and species-specific differences in innate activity can be explored. Lastly, techniques from efficient coding and neural modeling can be used across their respective disciplines to improve the current state of developmental V1 modeling.

The initial, monocular model was clearly limited in scope, but provided the foundation for later efforts. The binocular model is an obvious extension - particularly to demonstrate the role of binocular LGN/V1 activity as opposed to monocular retinal activity. However, in both models the patterns lack a temporal component. Chapter two reviewed how a topography of direction selective cells in the ferret only forms upon visual experience (Li et al. 2006). The unique way in which a temporal innate learning model will have to include temporal activity but not produce direction selectivity for a ferret model will be discussed in this chapter. Such a model would finish the obvious extensions of the initial model to include both binocular and spatiotemporal patterns.

Currently, the pattern parameters are chosen and sometimes constrained based on theoretical simplicity – for example along the percolation threshold, p_c. The constrained parameter space can be explored fully because the parameter space is limited. The models are evaluated by direct measurement of receptive fields either qualitatively, as in figure 4.2 by displaying how the produced filters appear like V1 gabor filters or quantitatively, as in figure 4.3 by demonstrating the distribution of

orientation bandwidths. The resulting filters can also be evaluated by their ability to efficiently encode natural signals; the goal of ICA is to find filters which maximize the likelihood of the data given the filters and the assumptions of ICA; this provides a fairly reasonable metric for efficiency. However, both evaluation methods are relatively indirect. In Chapter five, the binocular model shows that a perceptual judgment of depth (disparity) is possible given a set of binocular filters. Evolutionary algorithms would demonstrate how a particular species may optimize their ability to perform early depth discrimination given their species-specific environment. The possible differences between altricial and precocial visual young will be discussed.

Finally, as mentioned in the previous chapter, this dissertation ties together efficient coding and neural modeling. Although these levels of analysis concentrate on separate goals – e.g. receptive fields vs. topography – there are clear relations between the two. This is evident by the fact that the earliest innate learning approaches were neural models (Miikkulainen et al. 2005, Burger & Lang 1999), although these approaches suffer from the same issue as most neural models of development – a number of arbitrary assumptions because such detailed models are underconstrained by available physiological data. There has also been work overlapping these two analysis levels, but without the goal of understanding development. Topographic ICA (Hyvarinen & Hoyer 2001) showed how an ICA objective could be constrained to produce topographies, and Olshausen and colleagues (Rozell et al. 2008, Garrigues & Olshausen 2008) are demonstrating how sparse coding can be implemented in a physiologically plausible way. One goal in this respect can be to create a computational neural map model, similar to the LISSOM model (see Miikkulainen et al. 2005 for a review), which is an approximate instantiation of the efficient coding techniques mentioned previously. The goal of these future research efforts would be

to establish the link between normative efficient coding models and neural map models to better demonstrate the innate learning principle.

A spatiotemporal spontaneous activity model

Current models of wave activity, including the one used here, have a temporal component insofar as they all propagate activity, but patterns for analysis in this thesis were only chosen at wave termination. There are a number of reasons for this choice – both physiological and theoretical simplicity. The simplicity is clear - one fewer dimension (time) to encode, and fewer parameters to specify. Physiologically, certain temporal receptive field properties, like direction selectivity of V1 cells, are developed much later than orientation and ocular dominance, at least in animals like the ferret. There may also be a mismatch between speeds of spontaneous activity and real-world phenomena which suggest weaker support for the innate learning of temporal properties. However, the waves in retina, LGN, and V1 clearly have a temporal component – one of their most striking features – and it would be incomplete if it is not addressed in modeling.

The simplicity of the three-parameter monocular model was useful as an initial demonstration – only the fewest necessary parameters of the activity generation were included. Termination, as opposed to an arbitrary point in propagation, is a reasonable choice as the resulting patterns relate more closely to theoretical findings about percolation models. Nearly scale-invariant fractals are possible if the patterns are near the percolation threshold. This mathematical elegance permitted a modeling approach devoid of spatial or temporal scale – ideal for demonstrating the basic concept of innate learning. Also, by using spatial patterns, rather than spatiotemporal sequences,

the dimensionality of the analyzed patches was greatly reduced – a useful step for learning efficient codes.

The previous points of simplicity and efficiency were only for instructive or technical simplicity. The most significant reason to delay constructing a temporal activity model for later work is the unique developmental physiology of temporal coding in V1. In a number of animals direction selectivity is one of the topographically organized receptive field properties in V1. However, unlike ocular dominance or orientation selectivity, direction selectivity in ferrets is not present at eye-opening (Li et al. 2006). In fact, without visual experience in a critical period for two weeks after eye-opening, many other properties re-emerged, including orientation and spatial frequency, but not direction selectivity. Any temporal innate learning modeling approach will have to also address why direction selectivity may not be seen until visual experience. One possibility may be that the temporal properties of the spontaneous activity, which can propagate at relatively slow time scales, may not be appropriate for creating a direction selective code. Another possible explanation is that the temporal statistical properties of the activity may be adequate for innate learning, but a longer timescale for learning may be required. We know in the ferret that orientation column refinement and the number of orientation-selective cells improves after eye opening and can even improve in the first few weeks of eye opening in a dark-reared environment - although not as well as under normal visual experience (Chapman & Stryker 1993, White et al. 2001). Cortical columns and cells selective to a particular direction of motion are often selective to edges which are orthogonal to the direction of motion. For these two maps (orientation and direction of motion) to be in register, it may be simpler to form the orientation column map first and have direction selectivity form on top of it; this is suggested by the order of

development for these two features. In any case, the lack of direction of motion selectivity will have to be adequately addressed in any spatiotemporal model.

However, despite the concerns of increased complexity and physiological implications, there are two reasons for pursing a spatiotemporal model. First, the patterns, as they are currently generated, are inherently temporal as they are produced using an initiation-propagation-termination strategy. Also, although the dimensionality will increase substantially, the analysis technique can remain the same – e.g. instead of 16x16 pixel patches, they will be be 16x16x(# of frames). This is the benefit of an efficient coding approach – easy extensibility. On a more superficial level, the moving images of spontaneous activity, such as retinal waves, are particularly striking, and it is more satisfying to address this in a complete activity model.

Here we will take the opportunity to describe one approach to spatiotemporal modeling of spontaneous activity. First, given the lack of later endogenous activity characterization, a spatiotemporal model would inevitably relate to the retinal wave models previously mentioned (Butts et al. 1999, Godfrey & Swindale 2007) using a rate of independent, spontaneous firing for recruitable cells, fixed lengths of active firing, and refractory periods. An analysis of wave speed versus number of recruitable cells could be done, much like what was done in the work of Butts and colleagues for the fraction of recruitable amacrine cells. Also, a phase diagram showing what set of parameters generate large waves vs. small waves or fast vs. slow could demonstrate the qualitative aspects of the waves. From that point, temporal patch sequences could be taken and analyzed by ICA in the same way as the previous monocular and binocular models. To relate to the known physiology, the degree of direction

selectivity could be altered by affecting temporal properties of the activity to form orientation selective cells without necessarily introducing direction selectivity.

The spatiotemporal modeling technique can complete the approach of extending the dimensionality of the activity model. V1 cells form topographies of orientation, ocular dominance, spatial frequency, and direction selectivity; the first three dimensions have been addressed in this thesis. The unique development of direction selectivity, not present at eye-opening, can be included for a more complete visual development model. Spatiotemporal receptive fields are possible, and temporal properties, such as direction selectivity, can be analyzed. A spatiotemporal model is also the most palatable for demonstration of the innate learning paradigm, as the activity model is more straightforward to relate to the physiology of waves we view in recordings. With the success of the previous spatial monocular and binocular models, a spatiotemporal model will cap the end of the current approach to this problem. Specifically, the flexibility of the efficient coding innate learning paradigm will have been demonstrated by applying it to a significant number of diverse spontaneous activity models – monocular and binocular as well as spatial and spatiotemporal.

Evolutionary implications

The teleological purpose of innate learning is the evolutionary advantage it can provide. The simple genetic specification necessary for a simple method of pattern generation has particular evolutionary advantages that can be studied. There are also distinct differences in visual development between species that can only be addressed by noting the different evolutionary constraints – such as parental care vs. precocial young. It is reasonable to assume a tradeoff between the ability to learn flexibly and to learn quickly – analogous to the bias/variance tradeoff in machine learning. These

two advantages of innate learning can be best addressed through an evolutionary perspective – the faster rate of natural selection toward an improvement in fitness and the ability to address species-specific differences based on parental care.

One project can use an evolutionary paradigm to test and optimize depth judgment. From the presented binocular model it was shown (in figure 5.2) that a physiological distribution of disparity sensitive cells can be derived from an innate learning paradigm. As shown later in that chapter, activity of a set of receptive fields from the model can be combined to determine the disparity from a stereo pair of images. A given set of binocular filters can be tested by analyzing their ability to find the true depth map from a stereogram (as was demonstrated in figure 5.4). In the innate learning paradigm, spontaneous activity parameters can be chosen to maximize the system's ability to make depth discriminations. A possible approach using evolutionary algorithms would be to give a fitness penalty to individuals with receptive fields that do not properly reconstruct a given depth map. With this approach the competitive evolutionary advantage of innate learning would be readily apparent – a faster ability to process stereoscopic information at eye-opening.

However, an eventual optimal individual strategy as a product of evolution is not the only evolutionary advantage of innate learning - a faster rate of species improvement through natural selection provides a benefit to the species overall. Valsalam et al. (2007) demonstrated that innate patterns of activity can speed up model adaption using evolutionary algorithms for complex visual tasks - in their case for handwriting recognition, but the analogy can be considered for other complex visual tasks as well. Two different methods of biasing the visual recognition system prior to experience were compared; either the initial network weights were directly manipulated, or an

endogenous pattern generator was adapted to bias the weights prior to experience. Although directly adapting the initial network weights did not require an innate learning stage, such an approach was shown to be at a disadvantage; the endogenous learning approach provided the fastest improvement between generations. Although this work was demonstrated for handwriting recognition, the same approach may be applied in a more physiological setting to depth discrimination using the binocular model.

The second direction in an evolutionary context is to explore to what extent the animal should form a fully optimized V1 before eye-opening versus being plastic to changes in the environment and growth of the animal. Given a particular choice of spontaneous activity parameters, it is possible to evaluate the efficiency of the derived filters to natural scenes, but should one choose parameters to be most efficient for natural scenes at eye-opening or to minimize the time it takes to adapt to natural scenes upon a fixed amount of exposure to natural scenes? This is a fundamental question which changes what spontaneous activity patterns are considered optimal. Visually precocial young, like deer, require functioning visual systems quickly, while altricial young, like primates, are cared for after eye-opening and have a great deal of visual experience before there is selective pressure. The immature monocular properties of neural receptive fields in newborn primates (e.g. low spatial frequency) may be framed as an evolutionary choice to optimize under the additional assumption of a limited amount of natural experience. Optimizations under two different regimes can influence the properties of the learning before birth – functional visual codes based solely on innate experience versus those that can rely on a fixed amount of natural experience.

Ultimately, any general principle of sensory function must have an evolutionary advantage, otherwise it would not persist in the species. Biasing the visual system for behaviorally relevant processing objectives, like stereoscopic vision, is a clear advantage. However, even among the ways of biasing the visual system, there is an advantage of an innate learning approach versus directly adjusting initial connectivity. Precisely how this adaptation occurs depends upon the particular species. Visual system training could be optimized for use directly after eye-opening or after a fixed amount of natural experience. The two forms of optimality are on a sliding scale in the bias/variance tradeoff common to machine learning and represent the difference between visually precocial and altricial young. Because the innate learning paradigm is ultimately a product of evolution, the evolutionary paradigm will provide these useful insights in individual fitness, species rate of mutation, and species-specific optimization objectives.

Relating efficient coding and neural models of V1 development

Efficient coding models have the benefit of abstraction – simplicity that removes the details of implementation. However, efficient coding models cannot form a complete picture of the innate learning paradigm, as they typically cannot answer *how* the algorithms are biologically implemented. On the other hand, neural models of development form 2D maps and receptive fields without clear encoding objectives beyond reproducing the physiology. As stated previously, without a clear computational objective, these neural models can vary widely in their assumptions. The spontaneous activity of some of these models varies from uncorrelated noise to pixelated lines to gaussian blobs of activity, often with learning models of equivalently varying constraints. One of the goals of working on a computational neural model is to see precisely how these map models compare to efficient coding models. This

comparison will better relate a large body of computational work in adult vision to understanding the precise mechanisms of development in the brain. This direction consists of three sub-projects:

1. Apply topographic efficient coding models to spontaneous activity
2. Evaluate/adapt the receptive fields from the LISSOM neural map model using efficiency criteria.
3. Combine neurally-plausible sparse coding instantiations with LISSOM model assumptions.

First, one of the most interesting differences between the efficient encoding approaches to vision and the neural models of V1 formation is the issue of topography. The goal of V1 map-formation models is to build maps on which the selectivities of V1 cells to features vary characteristically, such as orientation and ocular dominance. The current V1 development model presented in this dissertation uses efficient coding to produce receptive fields but they are not arranged in any particular topography; neural filters inherently interact in the learning algorithm, but there is no predetermined 2D relationship about neighboring filters. However, there are efficient coding model variants which create a topographic arrangement. Hyvarinen and Hoyer (2001) applied a variant of independent components analysis, called topographic ICA, to natural images. Their approach produced a topographic map of neural filters which vary smoothly over position, orientation, and spatial frequency but randomly over gabor filter phase. The first approach can be to simply apply topographic ICA, or a more suitable variant, to patterns of spontaneous activity. This will demonstrate that efficient coding models of development can be evaluated in a similar way to the main goal of many neural models – V1 topography formation.

Second, not only can efficient coding models be more constrained to relate to neural map models topographically, but neural map models can be evaluated by standards of efficient coding. For example, most neural map models produce receptive fields that match those found in neurophysiology. Efficient coding techniques also produce receptive fields, but have the additional benefit of applying objective "fitness" functions – evaluating the sparseness or statistical independence of the code when presented with a set of stimuli. We can directly evaluate how efficient the code is from a neural map model. For example, one can use the LISSOM family of neural map models from Miikkulainen and colleagues (see Miikkulainen et al. 2005 for a review) and evaluate the resulting receptive fields for their sparseness or statistical independence upon presentation of natural scenes. Parameters of either the network learning models or the activity generation can be adapted to increase the efficiency of the resulting code. This work will relate the two levels of analysis by more than just a shared match to the physiology.

Third, one of the common criticisms of efficient coding approaches is that they are typically not implemented in a physiologically plausible way. This is in contrast to the fact that related neural implementations can not only be plausible, but also sometimes simpler. For example, principal components analysis (PCA) is performed in exceptionally nonphysiological ways (eigenvalue decomposition of a covariance matrix). However, a simple modification of Hebb's rule (called Oja's rule) can perform PCA on a data set in a neurally plausible way. There is recent work on finding ways in which sparse coding can be instantiated in neural architecture. In the traditional form, it relies on matrix calculations and implicit interactions within the entire population of neural filters. Recent approaches by Bruno Olshausen and

colleagues (Rozell et al. 2008, Garrigues & Olshausen 2008) show how sparse coding can be instantiated in the brain. Local neural circuits are used to instantiate a competitive, sparse coding mechanism, and horizontal connections arise due to the statistical properties of the images used for training. By contrast, local, horizontal interactions are *a priori* assumptions in the LISSOM family of models. The extent to which neural map models like LISSOM implement a sparse coding strategy and the effects of those differences can be further explored. Recent developments in both fields can help demonstrate the convergence of neural models and efficient coding, and thus result in an improved understanding of early visual development.

The future directions are intended to further demonstrate the inherent advantage in the innate learning approach. First, the spatiotemporal model is an extension made simpler by the use of an abstract, efficient coding learning method; simple adjustments to the activity model and use of the same learning algorithm can demonstrate interpretable results. The binocular model showed a perceptual interpretation (developing stereopsis) which can be exploited to show direct evolutionary significance. This evolutionary interpretation can show not only the faster individual adaptability at eye-opening, but also a faster trajectory for the species evolution. Finally, although neural-level modeling and efficient coding are only different levels of analysis, little work has been done linking the two approaches. Adding plausible constraints to efficient coding models or efficiency criteria to neural models has the potential to bring these two directions closer together, so work in both fields can be better informed. Ultimately, these research directions will better position disparate modeling approaches to more clearly validate the innate learning principle. These additional directions would also help us to better clarify some of the long-term objectives of this body of work, as will be discussed in the final chapter.

Chapter 8: Broader implications

In this chapter we explore the potential long-term benefits of an innate visual learning paradigm. The impact may reach across multiple domains including other sensory systems, high-level perception (e.g. faces), adult processing, and human health. To explain, V1 is not the only area which may benefit from an innate learning approach. The auditory and motor systems both require spontaneous activity during development, and there are early indications of a possible innate learning interpretation. Higher-level vision (e.g. face recognition) is often biased toward processing certain visual patterns, and this bias can occur through minor alterations in spontaneous activity. Although we discuss how the learning method is the same both before and after eye-opening, there is also work showing that spontaneous activity may also play a role in learning after visual experience. And finally, any basic research on human visual function has potential medical benefits, especially to deficits such as amblyopia where poor vision is a consequence of poor neural processing rather than poor optics. The potential applicability of innate learning outside of V1, development, or even basic research will demonstrate another reason for further pursuing this work.

The potential for innate learning in other sensorimotor systems

Endogenous neural activity occurs at the earliest stages of neural processing and is necessary for proper development of vision, audition, motor control, and other sensorimotor areas. In this section, we will briefly review the role of this activity in development prior to experience. The goal of this review is to relate what is known about spontaneous activity in these areas to our current understanding of similar activity in vision; although work has shown the necessity of spontaneous activity in all

these sensory areas, the functional consequence – such as a role in innate learning – requires a more thorough body of research. In a similar way to how retinal wave activity was reviewed to relate to possible innate learning in vision, low-level spontaneous activity in other sensory areas will be reviewed to consider the possibility of innate learning in other visual systems. This analogy is done with the intention of highlighting the similarities in the generation of these patterns and the developmental effects. For example, all these systems have a particular topography which is disrupted without proper activity-dependent innervation. In the visual and auditory systems, retinotopy and tonotopy are preserved with molecular guidance cues, but depend upon activity for refinement, and in olfaction the maintenance of a topography in the olfactory bulb depends critically upon activity.

Some of the earliest work on endogenous activity began in the motor system, with an awareness of embryonic motility for nearly a century. Later experiments showed this activity occurs well before sensory innervation (Bekoff et al. 1975). Also, there is a tight coupling between the somatosensory and motor systems for proprioception, and the physical movements caused by this activity aid this integration. More fundamentally, the motor system is performing the same translation between a physical world and an efficient, functional neural coding of that information, but only in the opposite direction of sensory systems. The way in which natural and repeated motions are encoded depends upon the morphology of the muscles and joints and the particular movements performed – this requires dynamic learning in the same way sensory systems universally adapt to natural inputs. Below we review the nature of auditory and motor spontaneous activity to understand the implication of the visual system work presented previously.

Auditory: Over a decade ago, spontaneous synchronous bursting activity was found in the auditory system. In chick embryos, spontaneous bursting has been measured in the auditory nerve (Lippe 1994). Sound thresholds in younger embryos during this activity are high (110+ dB), and activity was unaffected by removal of the middle ear ossicle, so the activity appeared to be endogenously generated within the cochlea. The role of auditory spontaneous activity has been studied in both mammals and birds, showing it's general presence across species (Friauf and Lohmann 1999).

Upon hearing a sound, the mechanical properties of the cochlea cause vibrations which are spatially separated by frequency and transduced by inner hair cells (IHCs); this tonotopic map is reflected in the rest of the auditory hierarchy and even into auditory cortex. Recent work in cats (Leake et al. 2006) has shown that prenatal and neonatal auditory neural activity is necessary for refining the topographic projections to the cochlear nucleus, and without this activity the nucleus was severely reduced in size. Recent work in rats (Tritsch et al. 2007) has shown the origin of this activity to be, in part, regulated by ATP release of nearby supportive cells. These cells, located in Kolliker's organ, have slow dynamics with rise times of the release being on the order of 1-2 seconds. The inward currents of inner hair cells had similar rise times occurring at a frequency of roughly once every 20 seconds. Activity is correlated highest among nearby IHCs but is statistically significant at distances as far as one octave (~250 m). Voltage clamped recording of support cells and IHCs showed each inward IHC current rise coincides with support cell activity. At the onset of hearing, these support cells stop firing and bursting, correlated spontaneous auditory activity shifts to the adult-like continuous discharge.

At this point it may be helpful to point out some of the similarities between visual and

auditory spontaneous activity. The auditory system uses spontaneous activity to refine its tonotopic projections in much the same way the visual system relies upon retinotopic activity. This activity occurs at the earliest transduction stages, in cells that hyperpolarize but don't emit action potentials. The developmental timing of this spontaneous activity is similar – occurring before birth but effectively ending at the onset of natural vision or audition. Interestingly the timescales of the particular patterns are particularly slow, with bursts of activity occurring a few times per minute and lasting on the order of one second. One reason for the presence of bursting across developmental systems is that it may be a necessary method for reliable information transmission in developing synapses (Lisman 1997). As an aside, correlated firing between the two ears must be integrated for proper localization using interaural time and intensity differences, which are known to occur in the auditory brainstem; although experimental evidence is lacking, spontaneous activity in later auditory stages may be necessary for integrating information from the two ears the same way LGN and V1 integrate binocular information.

The auditory system also presents another interesting analogy for the transition from innate to external sources of learning. Ferrets, mice, and cats are born with closed eye-lids, and for a short amount of time this leads to an amount of degraded visual experience through the eye-lids. As measured in ferrets, spontaneous waves of neural activity occur at this time, however, vision through closed eye-lids also stimulates V1 and helps the early visual system mature (Krug et al. 2001). Similarly in the auditory system, degraded auditory information can enter through the womb. For example, DeCasper and colleages (DeCasper & Fifer 1980, DeCasper & Spence 1986) demonstrated that children less than two days old preferred their mother's voice to another woman's (the same did not apply to fathers until weeks later). Also, mothers

were told to read *The Cat in the Hat* twice per day for the last 6.5 weeks of pregnancy, and the newborns preferred to hear a recording of their mothers reading that story compared to another story. However, this exposure to degraded sensory experience can vary significantly between species – humans are born being able to hear and with eyes that open (although also with poor optics) whereas many animals, such as cats, have a period of time with closed eyes and the ear canal is closed, where in both cases it is difficult to elicit neural responses (Walsh 1986).

Motor: The presence of spontaneous motility in developing animals is common knowledge. A human fetus moves in the womb early in life, but directly studying movements in the womb of any animal is difficult. Fortunately, the chick egg provides a readily accessible and controlled environment for embryonic growth and study, and has hence been the main model system for this research. Since neural models of this activity are less developed than retinal activity models, we will primarily discuss the overt effects of this activity and its disruption. The research which provided the basic delineation of movement patterns into three types in the chick embryo was done by visually observing embryonic motility through a hole in the shell (Hamburger and Oppenheim 1967). Already at embryonic day 3.5 (E3.5) type I movements are present as uncoordinated, random, jerky, small amplitude movements present until hatching. At E11 type II movements appear as sudden, rapid wriggles and startles of the whole body. At E17, type III movements are characterized as hatching and prehatching movements – tucking and piping of the egg shell and represent the first goal-directed, coordinated movements. Further behavioral measures have shown a more detailed, kinematic analysis of spontaneous movement (Bradley 1999). Early activity appeared to be more periodic and frequent, but less coordinated between joints. The transition from type I to type II was gradual; the amplitude of the movements increased while

they became less frequent, with a mixture of slower movements and rapid, abrupt movements. This is believed to be a modulatory effect of higher-level connections. Also, in the I-II transition, intralimb correlations increased while the interlimb correlations decreased. This, for example, may play a role in the functional differentiation of wing and leg movements.

Spontaneous motor activity is necessary for proper skeletal/muscular development. This has been shown by a decrease in skeletal muscle and bone mass if this early activity is silenced (Hall & Herring 1990). It has also been known for over half a century that this activity is necessary for proper patterning and development of the neuromuscular junction (Fatt & Katz 1952). Later work has shown how this activity is necessary for proper morphological, electrophysiological, and molecular development of spinal motor neurons (Kalb & Hockfield 1992). More recent evidence suggests that within spinal networks initial connections are guided by molecular cues, but proper patterning and refinement depends upon spontaneous activity (O'Donovan et al. 1998). However, we can discuss in a more general, functional sense why the effects of spontaneous movements may be beneficial in development.

In relating this work to other vertebrates, one must keep in mind that the functional requirements of chicks after hatching are much different than mammals at birth. It is more critical for chicks to develop coordinated movements like walking and running before hatching whereas human infants have time to develop these skills – making the timecourse of comparison different, but nonetheless informative. In the motor system there are a number of feedback-driven systems which must be in place hierarchically - systems for posture and muscle tone, muscle synergies, motion in 3D space, and organizing complex actions (Todorov 2004). At the level of the spinal cord, the

system must integrate proprioceptive and motor information dynamically and adaptively due to the constantly changing state of the animal during growth and beyond. Primary motor cortex activity produces a high-level coordination on particular muscles and joints in a way that integrates multiple joint mechanics and current physical states to calculate necessary muscle force. This early spontaneous spinal activity, especially before modulatory effects from higher-level inputs, may be necessary for the lower-level integration of sensorimotor information. The type I movements may help to establish certain muscle synergies. Proprioceptive integration necessary for flexible, complex actions in 3D may be integrated at this time for individual joints in type I motility, and coordination of entire limbs at the type II stage of embryonic development when intralimb coordination increases and modulatory effects from higher-level motor areas are present.

Summary for all modalities: There are many similarities between the three main modalities covered – vision, audition, and motor. In every case initial connectivity does not require neural activity, but is refined by it. Proper development requires balanced, competitive interactions for adequate representation, with activity dependence continuing into adulthood. This spontaneous activity dependence appears universal to sensorimotor systems as similar statements have been made for olfactory bulb patterning (Yu et al. 2004) and patterning of the somatosensory system (O'Leary et al. 1994; Erzurumlu and Kind 2001). This refinement/maintenance role of spontaneous activity is evident in the visual system in ocular dominance or orientation column formation where both have crude connectivity without spontaneous activity but are later refined by it. However, spontaneous activity is necessary for initial formation in some cases; for example, segregation of eye and ON/OFF-center cell layers in the LGN requires early activity (Shatz & Stryker 1988). The reasons for

which types of patterning require activity and which begin with axon guidance cues is not clear *a priori* but can often be determined experimentally by analyzing neural connections and functional topography at the relevant developmental timecourses.

The location and timing of this activity is conserved in each domain. In audition and vision, this occurs at the earliest locations of neural activity prior to spiking – the cholinergic amacrine cell in vision (Zhou 1998) and the inner hair cell in audition (the motor system cannot be any earlier than spontaneous evoked motion). Spontaneous activity appears to occur in the first location in which topographic information is available. For example, in olfaction similar receptors are randomly dispersed throughout the olfactory epithelium, and developmental spontaneous activity only occurs at the location of similar receptor aggregation in the olfactory bulb. In all senses, the developmental timing of activity begins soon after axon guidance cues form connections but effectively ends when the system is put to use. Ferret LGN/V1 spontaneous activity effectively ends before eye opening, auditory hair cell spontaneous activity ends before the onset of hearing, and spontaneous motor movement transitions from coordinated muscle movement to goal-related behavior – type II to type III spontaneous activity in the chick example. This progression of spontaneous activity up the sensorimotor hierarchy appears universal in the modalities mentioned. Motor activity begins with individual joints and increases correlations of muscle activity within limbs until goal-related behaviors emerge. In humans this exploratory, spontaneous motor movement occurs even into infancy where motor commands and proprioceptive sensation are combined with other sensory information. In vision, retinal activity patterns are uncorrelated between eyes and later LGN/V1 activity produces binocular correlations. The extent of binaural integration from spontaneous activity is yet to be measured experimentally, but may proceed similarly,

as the lateral and medial superior olive integrate binaural auditory signals for localization by interaural intensity and time differences, respectively. In each of these sensory systems, this activity occurs in the earliest topographic embedding of neural structures. With respect to timing, this occurs before active experience and proceeds through a sequence of lower to higher levels.

The main purpose of this section was not simply to draw analogies – it is to encourage a possible interpretation of activity in these areas as innate learning. The visual system has had the benefit of easily-manipulable inputs, work related to computation on those inputs, accessible neural processing (like the retina), and a history of related research. Since 1990, the progress from measuring retinal waves to our knowledge of innate learning has benefited from these advantages. The current majority of spontaneous activity research in other sensory modalities may be following a similar history to the initial reactions of interpreting retinal waves. Topology and refinement of neural connectivity are the primary focus of measurements and later stages of processing (like LGN/V1 activity) have received less attention. In order to pursue an innate learning interpretation for these stages, further experimental work is necessary.

The role of innate learning beyond V1: face perception

This dissertation has focused on early low-level processing, but the innate biasing of visual processing based on intrinsic patterns of activity may occur beyond V1. It is advantageous for young animals to be responsive, or at least hyper vigilant, to certain stable, highly-consequential visual cues. For example, macaques appear to have an innate tendency to learn fear associations for snakes compared to other related images (Mineka et al. 1984, Cook and Meneka, 1989). In humans, there appears to be an attentional bias even in young children to snakes compared to other nonthreatening

stimuli (Lobue and DeLoache 2008). Also, in humans there appears to be a specialization for processing faces in adults, with some researchers claiming a face-dedicated area in cortex (Kanwisher et al. 1997). Here we will assume a gradual development of these "specializations" by noting that they are likely to be caused by innate biases toward related features, rather than explicit specializations for snakes or faces; for example, the attentional bias for snakes may simply be a bias involving contoured, snake-like visual patterns. The rest of this section will expand upon the innate face bias and how it may relate to endogenously generated patterns. We will rely primarily on the experimental results from young infant experiments of Johnson and Morton (1991). The relation of the innate learning approach will be discussed in the context of their conceptual model of face preference (CONSPEC/CONLERN), and the endogenous pattern learning simulation of innate face preference of Bednar and Miikkulainen (2003).

Methodologically, Johnson and Morton had experimenters expose newborn infants to various configurations of facelike and non-facelike stimuli. Infant preference was measured by how long they track the stimuli across their visual field. The results of these experiments show a preference for configurations with parts representing eyes, noses, and mouths in an outline of a head. There are some exceptions, for example, checkerboard patterns presented within a head-shaped outline are preferred above their more facelike stimuli. Although there appears to be a preference for these artificial, schematic faces compared to real faces early on, after 5 months the preference for static, schematic faces is no longer present (Johnson and Morton 1991). It should be noted that throughout this time newborns are capable of learning individual faces (Slater 1993). These results can be interpreted as an innate orientation toward faces after birth, when learning occurs rather quickly. Precisely how this bias forms and

when the handoff is made to experience-dependent mechanisms is not clear from these studies.

The two models to be discussed here are Johnson and Morton's conceptual model and Bednar and Miikkulainen's computational model – an extension of the LISSOM computational neural map model of V1. Johnson and Morton (1991) proposed that infant face learning is performed by two hypothetical visual processing systems – CONSPEC and CONLERN. CONSPEC is fixed, assumed to be subcortical, and controls initial orientation toward facelike stimuli. CONLERN is plastic, cortical, and assumes control after approximately 6 weeks. As was previously mentioned, this approach is not entirely accurate, as Slater (1993) showed newborns were capable of recognizing individual faces, therefore both systems appear to be active at birth.

In contrast to the conceptual CONSPEC/CONLERN model, the HLISSOM model of Bednar and Miikkulainen is a computational model. Their computational model consists of a V1 layer, using their LISSOM computational V1 map formation algorithm, and a face selective area (FSA). The V1 map is formed by simple patterns of endogenous activity – in this case a series of randomly presented circles of fixed size. The FSA is trained by upright three-dot configurations at roughly the same scale as two eyes and a nose/mouth presented to an infant. The purported source of this type of activity is the ponto-geniculate waves present in REM sleep, however no precise relation to physiology was mentioned. Preference in this model is determined by which stimuli produce greater activity in FSA or V1 with more weight given to higher FSA activity.

Results from the HLISSOM model show an innate bias toward schematic upright faces, with some non-facelike stimuli preferences (like the checkerboard face pattern). Additionally, the model was later trained on natural images of faces and other objects – particularly relevant to an innate learning approach. The same model that was capable of an innate bias for faces was also capable of learning features of individual, presented images. There appears to be no conceptual need for a separation of fixed and plastic face preferences like CONSPEC and CONLERN. The same learning algorithm can be used on innate patterns as well as externally presented patterns, even when the goal is to bias the system toward high-level interpretations such as faces.

From these experiments and models, one can see that innate learning may also play a role in higher-level vision. Previous models with fixed and learned components can be replaced by models which use the same learning method both prior to and after visual experience. Although the innate three-dot patterns in the HLISSOM model seem rather artificial, it is possible that more naturalistic pattern generation techniques could be used. Methods like cellular automata can use fairly simple rules to generate patterns with certain consistencies, similar to how spots form on the wings of a butterfly. It is not unreasonable to assume the higher-level visual system, whether it is cortical or subcortical, could employ such a simple pattern generation technique to produce a critical, consequential attentional bias. In this way, innate learning may not only form early, low-level sensory maps like V1, but may also provide high-level attentional biases, like those toward snakes or faces.

The computational role of spontaneous activity after experience

The models presented in this thesis relied on the fact that similar learning can take place first on spontaneous activity and then later on natural experience after eye

opening; however, visual spontaneous activity does not end at eye-opening. The role of spontaneous activity before as well as after experience may provide another bridging point between these traditionally separate stages of visual development. In this section, we consider two separate types of adult spontaneous activity. First, a potential statistical relationship between spontaneous and evoked neural activity in V1 will be considered. The innate learning approach may be consistent with an interpretation of activity before eye-opening as a bayesian prior for natural scene statistics. At least in the adult, spontaneous activity and firing rate variability in evoked activity have been related from a bayesian, statistical point of view. The second discussion point about adult spontaneous activity is the hippocampus and its role during sleep and memory consolidation. Spontaneous activity in the hippocampus has been linked to adult experiences and is related to the process of dreaming. Adult spontaneous activity in V1 and the hippocampus will show that the innate learning principle may not be limited to development alone.

A bayesian view of adult V1 spontaneous activity: Spontaneous activity prior to experience may be related to neural variability found in adult neural responses. In this section, we will attempt to reformulate the innate learning interpretation of perinatal spontaneous activity to bayesian interpretations of adult neural response variability. The key insight relating the two is that activity prior to experience is a product of the collective experience of natural images in the species over evolutionary time, while adult neural variability is related to the individual experiences of the animal.

To better formalize the ideas of this section, we will refer to the statement of Bayes' theorem in equation 8.1 where 'x' represents the measured visual inputs and 's' is the inferred sources from the real world. The goal of the visual system is to infer the

correct sources given the data, that is, to calculate p(s|x).

(equation 8.1) $p(s|x) = p(x|s)\,p(s) / p(x) \propto p(x|s)\,p(s)$

In relation to V1 and sparse coding, the inferred sources are represented by the firing rates of neurons, and the prior, p(s), is that the firing rates come from a sparse distribution. However, most efficient coding approaches (e.g. Olshausen & Field 1996, Bell & Sejnowski 1997) learn a mapping from data to firing rates that is deterministic. This represents one point of p(s|x) instead of the full posterior distribution – similar to a maximum a posteriori (MAP) choice for p(s|x). It is unlikely that the full posterior is represented only as a MAP estimate, but how can the visual system represent the full posterior probabilistically? Hyvarinen and Hoyer (2003) claim this distribution can be represented by variations in neural firing rates, either time-varying or at the population level. In this formulation, neural variability represented by p(s|x) is simply a monte carlo sampling from the posterior distribution of possible firing rates. This insight relates to the innate learning approach if we note that activity prior to eye-opening is not from a naïve distribution, but rather has been informed by the accumulated experience of natural image statistics throughout species evolution – treating p(s) as p(s|x): an evidence-based prior where 'x' is accumulated over evolutionary timescales. Hyvarinen and Hoyer consider this activity primarily in the adult, where the variability is a reflection of learned information in an individual, whereas prior to eye-opening the innate learning approach considers this activity as an evolutionary way of guiding development. As was previously mentioned, activity prior to experience need not be a naïve prior, but rather reflects what has been genetically instilled in the animals over evolutionary timescales, and the role after eye-opening may simply be the influence of direct evidence from the individual animal. In both cases, spontaneous activity may be related to a monte carlo sampling over the posterior from accumulated experience with natural scenes, with the role prior to

experience being used for training while after experience there may be a more direct behaviorally-relevant role.

Recent work by Berkes et al. (2009) demonstrates a relationship between spontaneous activity in ferret V1 and the average evoked activity upon presentation of natural stimuli. Also, the changes in spatial and temporal correlations in V1 spontaneous activity and natural scene-evoked activity appear to match over the first two months of visual experience in ferrets (Fiser et al. 2004). Additionally, a recent optical imaging study by Han et al. (2008) showed that spontaneous activity in rats following repeated exposure to simple visual stimuli matched the evoked activity for more than six minutes after the exposure ended. Much of the activity even in evoked stimuli represents ongoing circuit dynamics which are only slightly altered by the immediate individual experience. This has been corroborated in fMRI studies as well. Fox et al. (2005) were able to enhance BOLD signals by subtracting the ROI (Region Of Interest) activity with the measured spontaneous BOLD activity in connected areas that were inconsequential to the experiment. Spontaneous activity in adult V1 is well known, but the extent to which it represents a probabilistic interpretation is unclear. Further research is necessary to establish a statistical interpretation of spontaneous adult activity in V1.

Hippocampus: Although the focus here has been on spontaneous activity in low-level sensorimotor systems, another important area shown to be spontaneously active during development is the hippocampus (Ben-Ari, Cherubini et al. 1989). Although vastly different in location and function than early sensorimotor systems there is classical evidence for a functional purpose of spontaneous activity in the adult hippocampus. To understand this we will provide some background. Loosely speaking, the

hippocampus appears critical for one-shot or episodic learning, but not for procedural/repeated/semantic learning. Results of hippocampal lesions in humans, such as the famous case study of HM (Scoville and Milner 1957) demonstrate this role. Without a hippocampus one can learn new motor skills or new facts when repeated, but not integrate new, unique information. Hippocampal lesions also lead to retrograde amnesia which is temporally graded – more recent memories are more likely to be affected. Based on these results and other work, the suggested role of the hippocampus is to help consolidate experiences in the rest of the brain (Squire and Alvarez 1995). The subjective experience of spontaneous activity to which we are most aware is dreaming – purely endogenous activity within the brain, although the activity is often related to identifiable events within our conscious experience. The hippocampus is active during sleep to help integrate these experiences in other areas of cortex, leading to less hippocampal dependency for particular memories over time. Evidence for this theory is particularly striking when looking at place-cell firing in rats, where the activity of place cells on a 2D neural map in the hippocampus follows the same temporal pattern in sleep as the rat followed when traversing the path earlier that day (Skaggs and McNaughton 1996). Evidence in humans also suggests that it is a form of adult spontaneous activity which is necessary for proper consolidation of memories (Stickgold 2005). The role of spontaneous activity in adult hippocampus appears to be a method of innate learning based on prior experience for the purpose in integrating new information.

In this section we have discussed two fundamental roles of spontaneous activity in the adult. First, and most clearly related to the previous V1 models, adult spontaneous activity may be thought of as a sampling of the posterior distribution of firing rates over natural images. The precise contribution to this sampling from an innate bias and

an individual bias learned by natural experience is unclear. In low-level areas like V1, spontaneous activity is correlated with average evoked activity. This lends a probabilistic interpretation to spontaneous activity not only prior to but also after experience. The hippocampus provides another analogy for understanding adult spontaneous activity by showing how cognitive processes, such as dreaming, may play a role in memory consolidation. Neural activity appears to "replay" events in order to integrate them in the cortical representation. Spontaneous activity after experience in V1 may have a similar role in consolidation. This dissertation has shown how the learning algorithm may be shared between stages of development before and after experience, and as reviewed in this section there is an apparent possibility that spontaneous activity may also bridge these stages as well.

Significance to human health

Although the goal of this dissertation is basic research, the role of understanding early visual system development has potential medical benefits. Here we briefly mention the way this work may better help us to understand amblyopia (poor vision due to abnormal visual experience), an impairment which affects 1-3% of the population (Webber & Wood 2005). It has already been discussed how high-level coding models can help constrain neural models to be more physiologically plausible and thus testable. With better neural-level models, we will not only be able to predict the effects of various experimental manipulations, but have a better understanding of accidental abnormalities which can occur in visual development. In other words, many of the experimental manipulations applied to understand V1 development have direct equivalents in human impairments. Amblyopia can be caused by strabismus (improper alignment of the eyes) which alters the natural input into the eyes. It is also known that amblyopia can be caused by prenatal factors. Such congenital processing

impairments may be a result of improper neural activity before birth, and this work may help link efforts to understand similar disorders of sensory processing that can occur perinatally. Such models ultimately provide further insights into how such abnormalities can develop and the mechanism by which current and future corrective therapies can operate.

Conclusion

The principle of innate learning - that the visual system can learn from endogenous neural activity in a similar way to how it can learn based on visual experience - is most naturally understood from a normative, efficient coding perspective; this perspective in development is unique to this dissertation. In neurophysiology, the permissive versus instructive role of this activity has been debated. Previous neural modeling approaches have shown that spontaneous activity can help form a visual code but have provided little detail about the nature of the spontaneous activity; this lack of clarity stems from the reliance on placing particular constraints in the neural implementation of the learning model rather than the statistics of the input. The efficient coding models in this work focus on the statistical properties of the activity, and avoid implementation details. By doing so, this allows a focus on the nature of the activity that would be necessary for a learning algorithm which could be used both before and after eye-opening. This work is an example of how abstraction and generalization can lead to better inference – the activity necessary for this approach to work is closer to known wave activity models than any of the neural models previously mentioned.

This work brings together a number of disparate areas of research. First, there is the interaction between multiple levels of analysis in development – neurophysiology, neural modeling, and efficient coding. This dissertation also strengthens the link

between two stages of development - before and after eye opening. The same learning algorithm may be present in the two separate stages, and in this chapter we discussed the idea that spontaneous activity plays a role in learning after eye-opening as well. An abstract activity model brought together knowledge of both retinal wave activity and properties unique to LGN/V1 spontaneous activity. The literature on V1 efficient coding, primarily in the context of adult vision, has also been applied to development. V1 codes learned by spontaneous activity can also be evaluated using perceptual judgments like depth perception through stereopsis. With a perceptual component, the evolutionary reasoning and justification of innate learning can be more clearly elucidated. The relations between the multiple levels of analysis, LGN/V1 vs. retinal activity, development vs. adult, physiology vs. perception, and individual adaptation vs. species evolution are all important in understanding and expanding upon the innate learning principle. Although many disparate areas are involved, the intention is to bridge between these areas to make an interpretation of the experimental findings simpler overall.

The goal of this work is not simply to piece together work among different academic areas. The hope is that a validation of this innate learning approach leads to greater insights beyond early visual coding. The same efficient coding algorithms have been used in the auditory system as they have been in vision, and similar patterns of spontaneous activity exist in other sensory modalities. Spontaneous activity may play a similar role in these other sensorimotor systems. The visual system must bias attention toward many different stimuli, and one way this can be done is by learning from endogenous patterns of activity beyond V1, as has been explored in the learning face preferences. This dissertation has explored the use of the same learning algorithm prior to and after experience, and spontaneous activity may serve a similar

role after experience as well. And as was mentioned in the last section, better understanding of the underlying biology has the potential to impact human health. A more thorough understanding of experimental manipulations may equivalently relate to accidental impairments like amblyopia. Further validating the innate learning principle has the potential to not only inform V1 formation, but also provide a roadmap for broader research pursuits like understanding sensorimotor development and perhaps even adult visual processing.

APPENDIX

A-1: Abridged derivation of the fastICA algorithm

The efficient coding algorithm used in this thesis was the fastICA algorithm, as summarized more thoroughly in the following paper:

> Hyvärinen, A., & Oja, E. (1997). A fast fixed-point algorithm for independent component analysis. *Neural Computation*, *9*(7), 1483-1492.

This algorithm was chosen for its speed of convergence – a fixed point algorithm with a cubic rate of convergence – and its similarity to other efficient coding techniques used to understand the computational role of V1 – sparse coding (Olshausen & Field 1996) and ICA (Bell & Sejnowski 1997). Specifics for the fastICA variant will be briefly summarized. The relation to other forms of ICA is best summarized by Hyvarinen and Oja (1997), while a sketch of the relationship between ICA and sparse coding follows in the next section of the appendix (A-2).

To begin, the data is initially preprocessed. 16x16 pixel patches are randomly taken from a set of images and are collected as column vectors of length 256 into a matrix, X_0. This preprocessing is used to lower the dimensionality of the problem and simplify analysis by removing noise, speeding up convergence, as well as significantly reducing the number of estimated parameters. This process is known as "whitening"; specifically, the pixel luminance values are decorrelated and set to unit variance. This can be done using principal component analysis (PCA), and is equivalent to using the eigenvalue decomposition of the covariance matrix, shown in equation A1.1, to create the transformed data matrix, X, shown in equation A1.2. The " ' " represents the fact

that only a subset of the eigenvectors and eigenvalues are selected from the decomposition.

(A1.1) $X_0 X_0^T = E^T D E$ (by eigenvalue decomposition)

(A1.2) $X = E' D'^{-1/2} E'^T X_0 = W X_0$

After the data is transformed, the goal of ICA is to linearly transform the data into maximally independent components. This linear transformation is shown in equation A1.3, where columns in S represent the component activations (equivalent to neural firing rates) and columns in A represent the basis vectors which correspond to the activations in the rows of S.

(A1.3) $X = A S$, or equivalently, $S = A^T X$

Because of the preprocessing of the data in X and activations in S are to be decorrelated and of unit variance, A is an orthogonal matrix ($A^T = A^{-1}$), so the component bases are equivalent to a filter description in this whitened space – that is, the columns of A (interpreted as bases) are equal to the rows of A^{-1} (interpreted as filters). For the original 16x16 pixel interpretation, the components must be transformed back to the original dimensionality which introduces a difference between the filter and basis descriptions. For a more thorough discussion of the distinction between bases and filters and how the whitening process used here affects their representation in the original 16x16 pixel dimensionality, see appendix A-3.

The goal in finding the orthogonal matrix, A, is to find vector directions, 'a', which project the data into components that are as statistically independent as possible. As arbitrary mixtures of random variables are more likely to be gaussian, the suggested solution in many ICA models is to find components with are the least gaussian. For a zero-mean random variable, 'y', a good proxy for this is the fourth-order cumulant, kurtosis, given in equation A1.4.

(A1.4) $\text{kurt}(y) = E[y^4] - 3\,(E[y^2])^2$

In this case, we are searching for vector directions, 'a', which maximize this kurtosis of the filter activations while maintaining a scale where $||a|| = 1$. This brings us to the following equation.

(A1.5) $\text{kurt}(a^T x) = E[(a^T x)^4] - 3\,E[(a^T x)^2]^2 = E[(a^T x)^4] - 3\,||a||^4$

To this we add the constraint that $||a|| = 1$ by a cost function F which penalizes deviations from $||a||=1$. The specific form of F is not important to the derivation.

(A1.6) $\text{objective}(a) = E[(a^T x)^4] - 3\,||a||^4 + F(||a||^2)$

The objective is to maximize this function. Gradient ascent algorithms can be found by differentiating equation A1.6 and iterating in the direction of the gradient, as shown in the following equation.

(A1.7) $\Delta a \propto x(t)\,(a(t)^T x(t))^3 - 3\,||a(t)||^2\,a(t) + F'(||a(t)||^2)\,a(t)$

However, convergence to an appropriate component can potentially be slow. The insight in fastICA is to find a more direct solution by solving for the fixed points – the points where the expectation of equation A1.7 is zero.

(A1.8) $E[x\,(a^T x)^3] - 3\,||a||^2\,a + F'(||a||^2)\,a = 0$

and this can be solved through a number of numerical algorithms. The following algorithm is the technique used in this thesis.

The steps of the fastICA (fixed-point ICA) algorithm:

1. initiate with a random initial vector a_0 with $||a_0||=1$
2. let $a_{k+1} = E[x(a_k^T x)^3] - 3\,a_k$. The expectation is estimated using a large sample of data vectors.
3. $a_{k+1} = a_{k+1} - A_{found}A_{found}^T\,a_{k+1}$ (A_{found} is constructed by previously found a's)
4. Renormalize a_{k+1}
5. go back to step 2 until $a_{k+1}^T a_k - 1$ is close enough to zero.

These five steps are run for each independent component. Step 2 is a solution method for equation A1.8. Step 3 removes the projection of 'a' on previously found IC's; A_{found} is a matrix where the column vectors are previously found 'a' vectors. Step 4 is a necessary renormalization, and step 5 is simply checking to see if the iteration converged to a solution. With a small data set, and cubic convergence, only 5-10 iterations may be needed to give a reasonable solution. Each component, 'a', is the equivalent of a filter in the whitened space, so to find the representation of the filter in the original space, 'a_0', we simply use the dewhiteneing matrix constructed by the eigenvalue decomposition in A1.1.

(A1.9) $$a_0 = E' D'^{1/2} E'^{T} a$$

This 'a_0' vector is in the original 16x16 pixel dimension, and can be viewed in the same way as an image patch for visual inspection.

A-2: Similarity of sparse and independent coding objectives

This section is a short mathematical description of why we equate the two efficient coding objectives, sparse and independent coding, for the purposes of this dissertation. We will begin with a statement of the independent coding objective, and derive how independence assumptions relate to sparse coding objectives.

For a linear code, encoding and decoding can be simply stated in matrix notation:

$$\mathbf{s} = \mathbf{A}^T \mathbf{x}$$

x – the image patch to be encoded – expressed as a column vector (see footnote[1])
A – the encoding matrix – columns represent linear filters [# of filters x dim of data]
(the transpose is to retain consistency of notation in the thesis)
s – the filter outputs

The goal is to maximize the statistical independence of the resulting filter responses (the s_i's). Note that mutual information of a set of random variables is 0 when they are independent and positive otherwise, so the problem can be restated as a minimization of the mutual information of the filter responses. Mutual information of a set of filters, I(s), is equivalent to the Kullback-Leibler divergence between $p(\mathbf{s})$ and $\prod_i p(s_i)$. Mathematically, this is also equivalent to the difference between the summed

[1] **Preprocessing – "whitening":** $\mathbf{x} = W \mathbf{x}_o$

Assume the original image patches, represented as column vectors, $\mathbf{x}_o$, have the mean luminance of the individual pixels over the set of image patches subtracted out. Although it is not necessary, it is possible to decorrelate the individual pixel luminance distributions and set the variances to 1 by transforming the data with a whitening matrix – found by using eigenvalue decomposition of the covariance matrix of the original data. Note that after obtaining the filters in A, finding the filter matrix

Covariance matrix, $C = E(\mathbf{x}_o \mathbf{x}_o^T)$ = by eigenvalue decomposition: $C = E D E^T$
We set, $W = E D^{-1/2} E^T$, so
$\mathbf{x} = W \mathbf{x}_o = E D^{-1/2} E^T \mathbf{x}_o$
Now the covariance matrix of the transformed $\mathbf{x}$ is I:
$E(\mathbf{x} \mathbf{x}^T) = E(W \mathbf{x}_o \mathbf{x}_o^T W^T) = W E(\mathbf{x}_o \mathbf{x}_o^T) W^T$
$= (E D^{-1/2} E^T) (E D E^T) (E D^{-1/2} E^T)^T$
$= E D^{-1/2} (E^T E) D (E^T E) D^{-1/2} E^T = E D^{-1/2} D D^{-1/2} E^T = E I E^T = E E^T = I$

entropy of the individual filter responses and the joint entropy. This is expressed below:

$$I(s) = \sum_i H(s_i) - H(\mathbf{s})$$

$$= \sum_i H(s_i) - H(A^T\mathbf{x})$$

$$= \sum_i H(s_i) - H(\mathbf{x}) - \log |\det(A^T)|$$

There are many free parameters in A^T for this minimization… most obviously, there is no scale on the rows/filters of A^T (scaling the filter response, s_i, would not affect the independence measure). To constrain this, we can arbitrarily fix the s_i at unit variance. Also, A^T can be constrained so the s_i are uncorrelated. These assumptions together lead to the following observation…

$$I = E(\mathbf{s}\,\mathbf{s}^T) = E(A^T\,\mathbf{x}\,\mathbf{x}^T A) = A^T E(\mathbf{x}\,\mathbf{x}^T) A\ldots$$

Note that if x is "whitened" data $E(\mathbf{x}\,\mathbf{x}^T) = I$, and we are looking for an orthogonal matrix where $A^T = A^{-1}$, which is a parameter estimation problem with half the free parameters as the original unconstrained problem. Taking the determinant of both sides yields…

$$1 = \det(I) = \det(A^T E(x\,x^T) A) = \det(A^T)\det(E(x\,x^T))\det(A) = \det(A)^2\det(E(x\,x^T))$$

$$\rightarrow \det(A) = \det(E(x\,x^T))^{-1/2}$$

This implies the constraints of uncorrelated filter outputs of unit variance leads to a det(A) which only depends upon the data (or 'I' if the data is whitened), not on the s_i, so it can be replaced by a constant for minimization.

$$I(s) = \sum_i H(s_i) - H(\mathbf{x}) - \log |\det(A)|$$

$$= \sum_i H(s_i) - H(\mathbf{x}) + C$$

Therefore, **to minimize I(s) we simply look for filters which minimize $H(s_i)$ with the covariance constraints on the s_i.** Now, calculating entropy explicitly is difficult – it requires a histogram-based calculation and it is not robust, especially in the

multidimensional case, so we're going to make some approximations and assumptions. According to information theory, the maximum entropy distribution for a fixed variance is the Gaussian distribution, therefore, in order to minimize $H(s_i)$ we try to find filters such that the s_i deviate maximally from Gaussian in terms of entropy. **This is equivalent to finding vectors in which the projections of the data are non-gaussian** (see footnote[2])

Intuitively, this makes sense. If we are looking for independent components, any random mixing of those components will, according to the central limit theorem, tend toward a gaussian distribution, and given no other prior information about the statistics of these distributions, this is the only assumption we can make about the data. Approximate measures for this deviation from gaussian rely on skewness and kurtosis calculations, related to the 3rd and 4th moments of a distribution, respectively. In the case of theoretical simple cells filters, they fire equally negatively or positively; the negative case implying the existence of a cell of opposite contrast since cells don't fire negatively. For this reason, theoretical simple cell filters are symmetric. Skewness is zero for symmetric distributions, leaving kurtosis as a widely used measure of how nongaussian the firing rate distribution is. Note that symmetric, highly kurtotic distributions will be spiked at the mean and have heavy tails. The goal in this case then is to find filters which usually have responses (s_i's) mostly near zero, but when they deviate from zero they are highly active.

To be clear, there are two relevant kurtosis measures for firing rates in a population of

[2] If we assume the statistics of x_i's are uncorrelated Gaussian random variables (which after whitening is not too unreasonable considering the CLT and the arbitrary mixing caused by the whitening matrix) the problem becomes: $I(\mathbf{s}) = C + \sum_i H(s_i) - H(x_{gauss}) = C + \sum_i [H(s_i) - H(x_{i,\,gauss})] = C + \sum_i [H(s_i) - H(\text{unit gaussian r.v.})] = C - \sum_i J(s_i)$, where J is a proxy for the entropy difference from a Gaussian distribution, where J = 0 if the distribution is Gaussian and J > 0 otherwise. The goal is to find filters with activities that are maximally non-gaussian in this J measure.

neural filters. Lifetime sparseness is a measure of how kurtotic the firing rate distribution is for individual filters upon presentation of many scenes (Willmore & Tolhurst 2001). This is the measure which is sought after in the previous derivation. However, the original definition of sparse coding is more related to population sparseness – the kurtosis of the firing rates of a population of filters when shown and individual image. This is equivalent to finding codes where only a few neurons are activity at any given time. Lehky et al. (2005) studied the relationship between these two measures (their terminology was selectivity and sparseness). A tight coupling between these two measures can be expected intuitively but in practice can depend on the data set. It was found for simple and complex cells that this tight coupling (termed ergodicity) is present in the cell responses to natural scenes. In other words, lifetime sparseness (selectivity) implies population sparseness (sparseness) for these cells.

In conclusion, by beginning with the goals of ICA and following a series of assumptions we find the objective of independence in practice is roughly equivalent to a sparse coding objective. That is why in this dissertation we do not attempt to disambiguate the two or justify the use of one objective over another.

A-3: Filters vs. bases, and the multiscale properties of ICA

This section focuses on the size and spatial frequency differences across the complete set of filters for encoding natural scenes and innate learning patterns; to address these concerns the relationship between filters and bases and its effect on the multiscale properties of the derived filters will be discussed in detail. First, the distinction between filters and bases will be made clear both mathematically and in relation to the physiology. Second, although the innate learning patterns can be formally scale invariant, there is a restriction in the scale of filters derived from the ICA algorithm; the main result is that the innate learning patterns produce filters and bases that are not at multiple spatial scales. A possibility for the innate learning of multiscale filters will be hypothesized; this will in part be addressed by a mixture model of two activity patterns, showing that mixtures of activity patterns can result in filters of multiple types, rather than an amalgamation of filter properties. Finally, this section will conclude with the implications of these results as they pertain to the physiology of multiple spatial scales in V1

Filters vs. bases: math and physiology

The ICA algorithm has been described in appendix section A-1. As with many other efficient coding approaches to understanding V1, it begins with "whitening" the inputs – that is, transforming the data to be uncorrelated and of unit variance. This is done by taking the data, in this case luminance values in square image patches, and constructing a matrix, X_0, where each pixel patch is represented as a separate column vector. The data matrix is decorrelated and set to unit variance using matrices from the eigenvalue decomposition. The data matrix in the lower-dimensional space (assuming only a subset of eigenvectors and eigenvalues are used) is now represented

as matrix X, and the transformations between the spaces are shown in equation A3.1 to A3.3. The " ' " designations are noting that only a subset of eigenvectors are being used, particularly important to the filters vs. bases distinction for ICA.

(A3.1) $X_0 X_0^T = E^T D E$ (by eigenvalue decomposition)

(A3.2) $X = E' D'^{-1/2} E'^T X_0 = W X_0$

(A3.3) $X'_0 = E' D'^{1/2} E'^T X = W^+ X$

After the data has been decorrelated and set to unit variance, ICA searches for the sources and the mixing matrix which satisfies the simple linear equation in A3.4. Each column of S represents the corresponding filter/neural firing/source activations for a column in the data vector X. In ICA formulations, the source activations are also forced to be uncorrelated and of unit variance. This forces the mixing matrix, A, to be an orthogonal matrix, and so by definition, $A^{-1} = A^T$, and the second part of A3.4 holds.

(A3.4) $X = A S, \quad S = A^T X$

What this means is that the columns of A correspond to the filters of the data in the whitened space. These same vectors also form the basis set in the whitened space because the mixing matrix is an orthogonal matrix.

The distinction between filters and bases become more prominent when projecting back to the original space. In most of the thesis, the original 16x16 pixel patches (256 dimensional) are whitened and reduced to a 100 dimensional space. The whitening matrix, W, in equation A3.2 is a 100 x 256 matrix, clearly not invertible. This causes a distinct difference when transforming filters and bases back to the original unwhitened space. The dewhitening matrix, W^+, in equation A3.3, a matrix constructed from the PCA components, is incidentally the Moore-Penrose inverse (the "pseudo inverse") of W. Although the filters and bases in the whitened space are equivalent, there are distinctions in the original space because of the whitening and dewhitening process.

Equation A3.5 shows that filters are represented as the rows of $A^T W$, or equivalently by transposition, $W^T a$, where 'a' is a column of the mixing matrix. Equation A3.6 shows the basis set in the original space, as represented by columns in W^+A, or simply W^+a.

(A3.5) $S = A^T X = (A^T W) X_0 \rightarrow$ filter in original space = $W^T a$

(A3.6) $X = A S \rightarrow X'_0 = W^+ X = (W^+ A) S \rightarrow$ basis in orig. space = $W^+ a$

Thus, the distinction between a filter and a basis (which are the same columns of A in the whitened space) comes from the transformation of that vector to the original space by W^T for filters or W^+ for bases.

What does this mean in more intuitive, physiological terms? Filters represent the transformation from data, or images, to neural responses. This is equivalent to the receptive field description in terms of physiology. The basis is the expectation of what is present within the receptive field. For example, if one wants to efficiently represent 1D signals that are simple square waves, a reasonable filter would be a difference filter to detect the jump in signal, but the basis would indicate values which extend beyond the boundary points. Simply put, the filter represents the receptive field, and the basis represents what the visual system assumes is under the receptive field when the neuron fires. This distinction can be seen in Figure A.1 at the end of this section. Similar to the 1D step function example given above, two-tone images can be represented well by localized difference filters, while the bases indicate a luminance difference between the two sides of that filter.

Scale invariant patterns do not necessarily produce filters at multiple scales

One of the interesting properties of the abstract spontaneous activity patterns as mentioned in chapter 4 is that at the percolation threshold the patterns are scale

invariant. That is, percolation networks at their critical point have the same statistical regularities regardless of how often you zoom out of the image. Of course, physiological patterns are not expected to be perfectly scale invariant, but there is a benefit to having emergent large-scale statistical properties from small-scale local interactions. Despite the relative scale invariance of the two-tone spontaneous activity patterns or thresholded noise patterns, a quick observation from viewing figure A.1 shows that the produced filters are present at only one scale/spatial frequency. One reason for single-scale filters being derived from nearly scale-invariant images is that the filter size is constrained by the size of the image patches as well as the dimensionality and fixed number of filters. However, this does not explain why all filters are at the same spatial scale. Looking at the filters produced from an efficient coding of natural images in figure A.2, we can see that natural images produce some filters of low spatial frequency, with an larger assortment of high spatial frequency filters. Clearly, the parameters of the efficient coding algorithm do not constrict the filters to be of only one scale.

The single-scale receptive fields in the 2-tone patterns may also be from the reduced need to represent anything other than a localized difference operation. As can be seen in the bases, the filters indicate the localized presence of a luminance edge. A series of these difference can be enough to encode an entire patch, as the images are relatively simple 2-tone images with edges of the same fractal dimension. The natural images, on the other hand, have additional structure beyond a simple edge which needs to be encoded – therefore prompting use of multiple spatial scales.

Throughout this thesis, activity patterns were constructed with only one set of percolation parameters. However, it is known that in the physiology of retinal waves,

the physiological equivalent to the parameter 'p' - the fraction of recruitable amacrine cells - varies over space and time. Therefore, the parameters of the activity patterns in any location can vary over time. To simulate this, and relate to the previous lack of multiscale filters, image patches from two different sets of 2-tone images (as were shown in figure 4.4) were equally mixed. Relatively straight edges from an amplitude exponent of -4, which produced elongated filters, were mixed with high fractal-dimension edges of exponent -1.6, which produced more localized, point-like receptive fields (filters were also shown in figure 4.4). The complete efficient coding results of that mixture are shown in figure A.3. Although the distinction is subtle in comparison to the filters from either set individually, the complete set appears to be a heterogeneous mixture of filters. Although this mixture mainly concerns the width of receptive fields, it is not unreasonable to assume that similar mixed manipulations could produce a heterogenous mixture of other filter properties, such as distinct spatial frequencies.

The presence of multiple scales at birth is not readily measurable as the spatial frequency of filters for macaques and ferrets is relatively low. Recall that previous research showed phase disparity and ocular dominance values near adult level in macaques (Chino et al. 1997, Maruko et al. 2008) but the monocular properties of low spatial frequency made measurement of these properties more difficult. Spatial frequency may be a selectivity that is less critical to establish early on, especially in altricial animals, shown in part by its lack of development in early vision. However, the efficient coding of natural images (e.g. figure A.2), as well as adult physiology, shows that filters at multiple scales should eventually be present. Whether a

mechanism like a mixture model can be used to bias multiscale development or the visual system waits until natural experience is unclear. Further research is necessary to establish the role of multiple spatial scales in early visual system development.

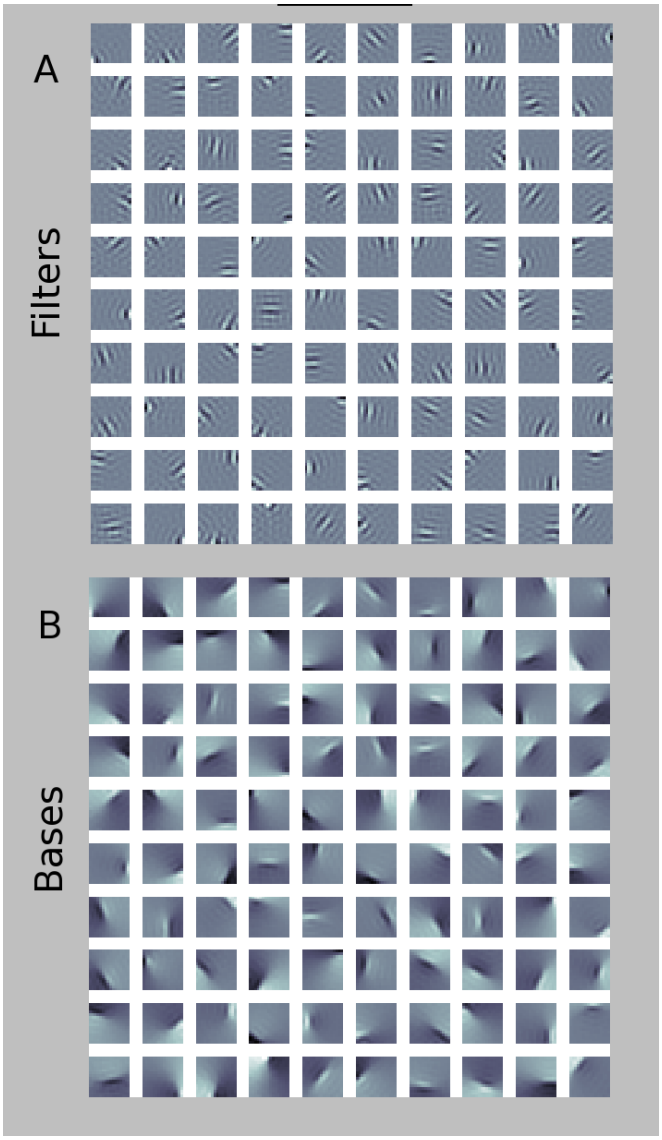

Figure A.1: All filters and bases derived from a representative 2-tone innate learning pattern

Presented are all 100 filters and their corresponding bases derived from an efficient coding of a pattern resembling spontaneous activity from chapter 4. In this case, the patterns were two-tone thresholded noise patterns with an amplitude exponent of -2.2, examples of which are shown in figure 4.4. ICA parameters were as outlined in the methods section of chapter 4. Note the same scale for all filters, and the strong low-frequency component found in the bases representation.

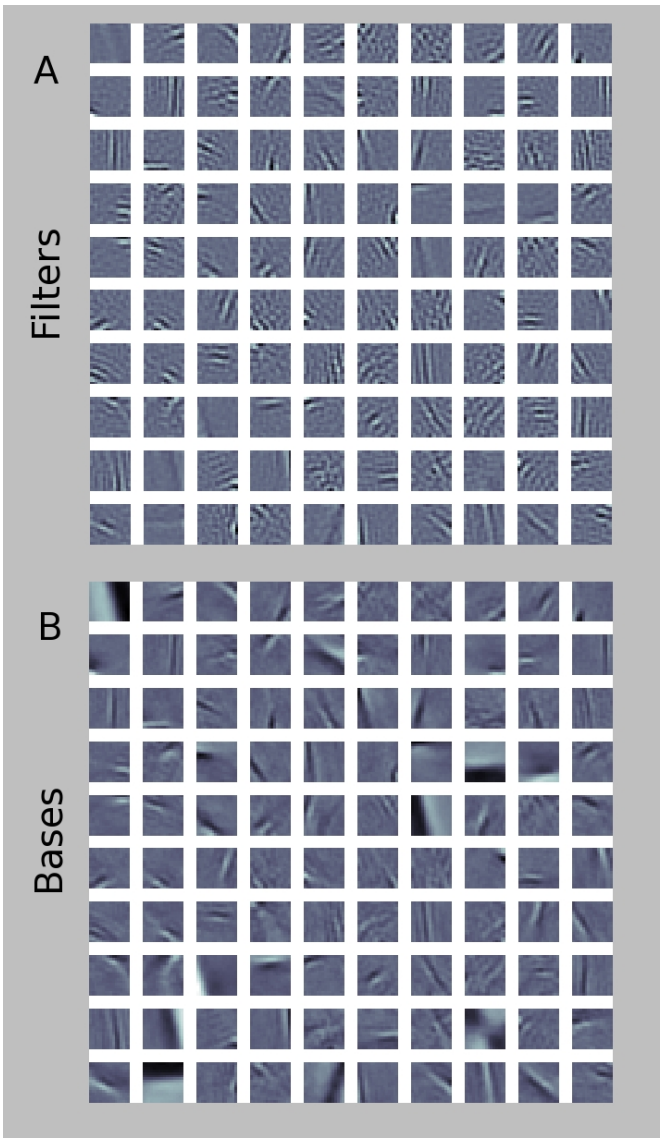

Figure A.2: Filters and bases derived from natural scenes

Presented are all 100 filters and their corresponding bases derived from an efficient coding of natural scenes. ICA parameters were outlined in the methods section of chapter 4. Note that there are filters at multiple spatial scales, more evident in the basis representation.

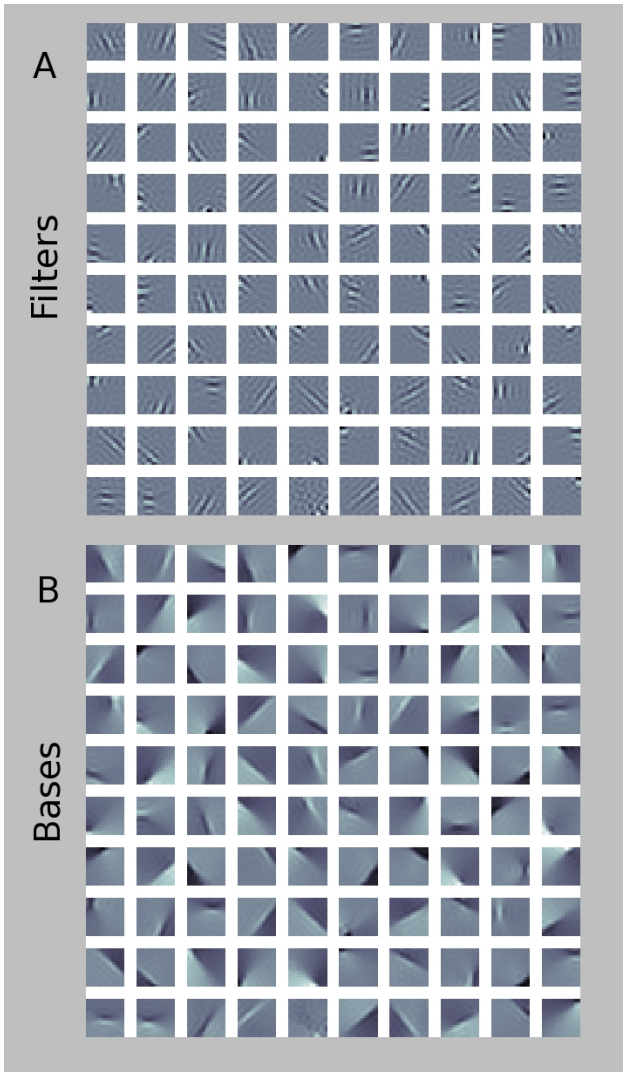

Figure A.3: Filters and bases derived from a mixed set of innate learning patterns

Presented are all 100 filters and their corresponding bases derived from an efficient coding of a mixture of patterns used in chapter 4. In this case, the images patches came equally from two different sets of patterns: two-tone thresholded noise patterns with an amplitude exponent of -1.6 and -4, as shown in figure 4.4. These two patterns represent high fractal dimension (1.49) and low fractal dimension (1.19) edges, producing point-like and elongated filters, respectively. ICA parameters were as outlined in the methods section of chapter 4. Note the heterogenous mixture of filter properties as shown in figure 4.4, rather than a homogenous averaging across all filters.

A-4: Source material

Portions of this thesis have been previously published, and sections have been taken from previous unpublished work – all originally written my Mark V. Albert. Here is a list of those sources, for full disclosure.

- Albert, M.V., Field, D.J. (in preparation) A binocular model of innate learning through spontaneous activity.
- Albert, M.V., Field, D.J. (submitted, neural computation) The possible role of LGN/V1 spontaneous activity as a training pattern for visual development
- Albert, M.V., Schnabel, A., Field, D.J. (2008) Innate Visual Learning through Spontaneous Activity Patterns. PLoS Computational Biology 4(8)
- Albert, M.V., Field, D.J. (2009) Neural Coding/Representation. Encyclopedia of Perception, Ed. E.B. Goldstein et al., SAGE Press.

Portions of A-exam qualifying papers on the following topics

- The functional and computational role of visual system nonlinearities
- Efficient coding in vision and audition
- The role of spontaneous activity in sensory and motor systems

REFERENCES

Adelson, E. H., & Bergen, J. R. (1985). Spatiotemporal energy models for the perception of motion. Journal of the Optical Society of America. A, Optics and Image Science, 2(2), 284-299.

Aharony, A., & Stauffer, D. (1994). Introduction to Percolation Theory (2nd ed.). Taylor & Francis.

Akerman, C. J., Smyth, D., & Thompson, I. D. (2002). Visual Experience before Eye-Opening and the Development of the Retinogeniculate Pathway. Neuron, 36(5), 869-879. doi: 10.1016/S0896-6273(02)01010-3.

Albert, M. V., Schnabel, A., & Field, D. J. (2008). Innate Visual Learning through Spontaneous Activity Patterns. PLoS Comput Biol, 4(8), e1000137. doi: 10.1371/journal.pcbi.1000137.

Albrecht, D. G., Geisler, W. S., & Crane, A. M. (2003). Nonlinear Properties of Visual Cortex Neurons: Temporal Dynamics, Stimulus Selectivity, Neural Performance. Chapter in The Visual Neurosciences, edited by L. Chalupa and J. Werner; MIT Press 2003.

Atick, J. J., & Redlich, A. N. (1992). What does the retina know about natural scenes? Neural Comput., 4(2), 196-210.

Attneave, F. (1954). Some informational aspects of visual perception. Psychological Review, 61(3), 183-193.

Attwell, D., & Laughlin, S. B. (2001). An energy budget for signaling in the grey matter of the brain. Journal of Cerebral Blood Flow and Metabolism, 21(10), 1133-1145. doi: 10.1097/00004647-200110000-00001.

Barlow, H. B. (1961). Possible principles underlying the transformation of sensory messages. WA Rosenblith (Ed.), Sensory communications. Cambridge, MA: MIT Press.

Barrow, H. G., Bray, A. J., & Budd, J. M. L. (1996). A self-organizing model of "color blob" formation. Neural Comput., 8(7), 1427-1448.

Bear, M. F. (1995). Critical periods in visual system development. Neuroscience in Medicine, 465-483.

Bednar, J., & Miikkulainen, R. (2003). Self-organization of spatiotemporal receptive fields and laterally connected direction and orientation maps. Neurocomputing, 52-54,

473-480.

Bednar, J. A., & Miikkulainen, R. (2003). Learning innate face preferences. Neural Comput., 15(7), 1525-1557.

Bednar, J. A., & Miikkulainen, R. (2004). Prenatal and postnatal development of laterally connected orientation maps. Neurocomputing, 58-60, 985-992. doi: 10.1016/j.neucom.2004.01.156.

Bednar, J. A., & Miikkulainen, R. (2006). Joint maps for orientation, eye, and direction preference in a self-organizing model of V1. Neurocomputing, 69(10-12), 1272-1276. doi: 10.1016/j.neucom.2005.12.090.

Bednar, J. A. (2002). Learning to see: genetic and environmental influences on visual development. Thesis from the Department of Computer Scienes, The University of Texas at Austin.

Bekoff, A., Stein, P. S., & Hamburger, V. (1975). Coordinated motor output in the hindlimb of the 7-day chick embryo. Proceedings of the National Academy of Sciences of the United States of America, 72(4), 1245–1248.

Ben-Ari, Y., Cherubini, E., Corradetti, R., & Gaiarsa, J. L. (1989). Giant synaptic potentials in immature rat CA3 hippocampal neurones. The Journal of Physiology, 416, 303-25.

Berkes, P., Orban, G., Lengyel, M., & Fiser, J. (2009). Matching spontaneous and evoked activity in V1: a hallmark of probabilistic inference. Computational and Systems Neuroscience (CoSyNe) conference, Salt Lake City, UT.

Berman, N., & Daw, N. W. (1977). Comparison of the critical periods for monocular and directional deprivation in cats. The Journal of Physiology, 265(1), 249–259.

Berns, G. S., Dayan, P., & Sejnowski, T. J. (1993). A Correlational Model for the Development of Disparity Selectivity in Visual Cortex that Depends on Prenatal and Postnatal Phases. Proceedings of the National Academy of Sciences of the United States of America, 90(17), 8277-8281.

Bradley, N. S. (1999). Transformations in Embryonic Motility in Chick: Kinematic Correlates of Type I and II Motility at E9 and E12. J Neurophysiol, 81(4), 1486-1494.

Burger, T., & Lang, E. W. (1999). An incremental Hebbian learning model of the primary visual cortex with lateral plasticity and real input patterns. Zeitschrift Für Naturforschung. C, Journal of Biosciences, 54(1-2), 128-40.

Burgi, P., & Grzywacz, N. (1994). Model for the pharmacological basis of spontaneous synchronous activity in developing retinas. J. Neurosci., 14(12), 7426-7439.

Butts, D. A., Feller, M. B., Shatz, C. J., & Rokhsar, D. S. (1999). Retinal waves are governed by collective network properties. J. Neuroscience, 19(9), 3580-93.

Cang, J., Rentería, R. C., Kaneko, M., Liu, X., Copenhagen, D. R., & Stryker, M. P. (2005). Development of precise maps in visual cortex requires patterned spontaneous activity in the retina. Neuron, 48(5), 797-809. doi: 10.1016/j.neuron.2005.09.015.

Carreira-Perpiñán, M. A., Lister, R. J., & Goodhill, G. J. (2005). A computational model for the development of multiple maps in primary visual cortex. Cerebral Cortex (New York, N.Y.: 1991), 15(8), 1222-1233. doi: 10.1093/cercor/bhi004.

Catsicas, M., Bonness, V., Becker, D., & Mobbs, P. (1998). Spontaneous Ca2+ transients and their transmission in the developing chick retina. Current Biology: CB, 8(5), 283-286.

Chandrasekaran, A. R., Plas, D. T., Gonzalez, E., & Crair, M. C. (2005). Evidence for an Instructive Role of Retinal Activity in Retinotopic Map Refinement in the Superior Colliculus of the Mouse. J. Neurosci., 25(29), 6929-6938. doi: 10.1523/JNEUROSCI.1470-05.2005.

Chapman, B., & Stryker, M. (1993). Development of orientation selectivity in ferret visual cortex and effects of deprivation. J. Neurosci., 13(12), 5251-5262.

Chapman, B., & Godecke, I. (2000). Cortical Cell Orientation Selectivity Fails to Develop in the Absence of ON-Center Retinal Ganglion Cell Activity. J. Neurosci., 20(5), 1922-1930.

Chapman, B., Jacobson, M. D., Reiter, H. O., & Stryker, M. P. (1986). Ocular dominance shift in kitten visual cortex caused by imbalance in retinal electrical activity. Nature, 324(6093), 154-156. doi: 10.1038/324154a0.

Chino, Y. M., Smith, E. L., Hatta, S., & Cheng, H. (1997). Postnatal development of binocular disparity sensitivity in neurons of the primate visual cortex. J. Neuroscience 17(1), 296-307.

Chiu, C., & Weliky, M. (2001). Spontaneous activity in developing ferret visual cortex in vivo. J. Neuroscience 21(22), 8906-14.

Chiu, C., & Weliky, M. (2002). Relationship of Correlated Spontaneous Activity to Functional Ocular Dominance Columns in the Developing Visual Cortex. Neuron, 35(6), 1123-1134. doi: 10.1016/S0896-6273(02)00867-X.

Ciuffreda, K. J., Levi, D. M., & Selenow, A. (1991). Amblyopia: Basic and Clinical Aspects. Butterworth-Heinemann.

Cook, M., & Mineka, S. (1989). Observational conditioning of fear to fear-relevant versus fear-irrelevant stimuli in rhesus monkeys. Journal of Abnormal Psychology, 98(4), 448-459.

Crair, M. C., Gillespie, D. C., & Stryker, M. P. (1998). The role of visual experience in the development of columns in cat visual cortex. Science, 279(5350), 566-570.

Crair, M. C., Horton, J. C., Antonini, A., & Stryker, M. P. (2001). Emergence of ocular dominance columns in cat visual cortex by 2 weeks of age. The Journal of Comparative Neurology, 430(2), 235-249.

Crowley, J. C., & Katz, L. C. (2000). Early Development of Ocular Dominance Columns. Science, 290(5495), 1321-1324. doi: 10.1126/science.290.5495.1321.

Cumming, B. G. (2002). An unexpected specialization for horizontal disparity in primate primary visual cortex. Nature, 418(6898), 633-636. doi: 10.1038/nature00909.

Daugman, J. (1980). Two-dimensional spectral analysis of cortical receptive field profiles. Vision Research, 20(10), 856, 847.

Daw, N. W., & Wyatt, H. J. (1976). Kittens reared in a unidirectional environment: evidence for a critical period. The Journal of Physiology, 257(1), 155–170.

De Valois, R. L., Yund, E. W., & Hepler, N. (1982). The orientation and direction selectivity of cells in macaque visual cortex. Vision Research, 22(5), 531-544.

DeCasper, A. J., & Spence, M. J. (1987). Prenatal maternal speech influences newborn's perception of speech sounds. Annual progress in child psychiatry and child development, 20, 5–25.

DeCasper, A., & Fifer, W. (1980). Of human bonding: newborns prefer their mothers' voices. Science, 208(4448), 1174-1176. doi: 10.1126/science.7375928.

Elstrott, J., & Feller, M. B. (2009). Vision and the establishment of direction-selectivity: a tale of two circuits. Current Opinion in Neurobiology. doi: 10.1016/j.conb.2009.03.004.

Erwin, E., & Schulten, K. (1995). Models of orientation and ocular dominance columns in the visual cortex: A critical comparison. Neural Computation, 7, 425-468. doi: 10.1.1.45.4146.

Erzurumlu, R. S., & Kind, P. C. (2001). Neural activity: sculptor of 'barrels' in the neocortex. Trends in Neurosciences, 24(10), 589-595.

Falconbridge, M. S., Stamps, R. L., & Badcock, D. R. (2006). A simple Hebbian/anti-Hebbian network learns the sparse, independent components of natural images. Neural Computation, 18(2), 415-429. doi: 10.1162/089976606775093891.

Fatt, P., & Katz, B. (1952). Spontaneous subthreshold activity at motor nerve endings. The Journal of Physiology, 117(1), 109–128.

Feller, M. B., Butts, D. A., Aaron, H. L., Rokhsar, D. S., & Shatz, C. J. (1997). Dynamic processes shape spatiotemporal properties of retinal waves. Neuron, 19(2), 293-306.

Feller, M. B., Wellis, D. P., Stellwagen, D., Werblin, F. S., & Shatz, C. J. (1996). Requirement for cholinergic synaptic transmission in the propagation of spontaneous retinal waves. Science, 272(5265), 1182-7.

Field, D. J. (1987). Relations between the statistics of natural images and the response properties of cortical cells. Journal of the Optical Society of America. A, Optics and Image Science, 4(12), 2379-2394.

Field, D. J. (1994). What is the goal of sensory coding? Neural Comput., 6(4), 559-601.

Field, D. J., & Tolhurst, D. J. (1986). The structure and symmetry of simple-cell receptive-field profiles in the cat's visual cortex. Proceedings of the Royal Society of London. Series B, Containing Papers of a Biological Character. Royal Society (Great Britain), 228(1253), 379-400.

Fiser, J., Chiu, C., & Weliky, M. (2004). Small modulation of ongoing cortical dynamics by sensory input during natural vision. Nature, 431(7008), 573-578. doi: 10.1038/nature02907.

Fox, M. D., Snyder, A. Z., Zacks, J. M., & Raichle, M. E. (2006). Coherent spontaneous activity accounts for trial-to-trial variability in human evoked brain responses. Nat Neurosci, 9(1), 23-25. doi: 10.1038/nn1616.

Freeman, R., & Ohzawa, I. (1992). Development of binocular vision in the kitten's striate cortex. J. Neurosci., 12(12), 4721-4736.

Frégnac, Y., & Imbert, M. (1984). Development of neuronal selectivity in primary visual cortex of cat. Physiological Reviews, 64(1), 325-434.

Friauf, E., & Lohmann, C. (1999). Development of auditory brainstem circuitry. Cell and Tissue Research, 297(2), 187-195. doi: 10.1007/s004410051346 .

Galli, L., & Maffei, L. (1988). Spontaneous impulse activity of rat retinal ganglion cells in prenatal life. Science, 242(4875), 90-91.

Garaschuk, O., Hanse, E., & Konnerth, A. (1998). Developmental profile and synaptic origin of early network oscillations in the CA1 region of rat neonatal hippocampus. The Journal of Physiology, 507 (Pt 1), 219-36.

Garrigues, P., & Olshausen, B. (2008). Learning Horizontal Connections in a Sparse Coding Model of V1. Advances in Neural Information Processing Systems.

Geman, S., Bienenstock, E., & Doursat, R. (1992). Neural Networks and the Bias/Variance Dilemma. Neural Computation, 4(1), 1-58. doi: 10.1162/neco.1992.4.1.1.

Godfrey, K. B., & Swindale, N. V. (2007). Retinal wave behavior through activity-dependent refractory periods. PLoS Computational Biology, 3(11), e245. doi: 10.1371/journal.pcbi.0030245.

Gödecke, I., Kim, D. S., Bonhoeffer, T., & Singer, W. (1997). Development of orientation preference maps in area 18 of kitten visual cortex. The European Journal of Neuroscience, 9(8), 1754-1762.

Grabska-Barwinska, A., & von der Malsburg, C. (2008). Establishment of a Scaffold for Orientation Maps in Primary Visual Cortex of Higher Mammals. J. Neurosci., 28(1), 249-257. doi: 10.1523/JNEUROSCI.5514-06.2008.

Graham, D. J., Chandler, D. M., & Field, D. J. (2006). Can the theory of "whitening" explain the center-surround properties of retinal ganglion cell receptive fields? Vision research, 46(18), 2901–2913. doi: 10.1016/j.visres.2006.03.008.

Hall, B. K., & Herring, S. W. (1990). Paralysis and growth of the musculoskeletal system in the embryonic chick. Journal of Morphology, 206(1), 45-56. doi: 10.1002/jmor.1052060105.

Hamburger, V., & Oppenheim, R. (1967). Prehatching motility and hatching behavior in the chick. The Journal of Experimental Zoology, 166(2), 171-203. doi: 10.1002/jez.1401660203.

Han, F., Caporale, N., & Dan, Y. (2008). Reverberation of recent visual experience in spontaneous cortical waves. Neuron, 60(2), 321-327. doi: 10.1016/j.neuron.2008.08.026.

Hopfield, J. J. (1982). Neural networks and physical systems with emergent collective computational abilities. Proceedings of the National Academy of Sciences of the United States of America, 79(8), 2554–2558.

Horton, J., & Hocking, D. (1996). An adult-like pattern of ocular dominance columns in striate cortex of newborn monkeys prior to visual experience. J. Neurosci., 16(5), 1791-1807.

Hoyer, P. O., & Hyvärinen, A. (2000). Independent component analysis applied to feature extraction from colour and stereo images. Network, 11(3), 191-210.

Hoyer, P. O., & Hyvarinen, A. (2003). Interpreting neural response variability as

Monte Carlo sampling of the posterior (p. 293). Advances in Neural Information Processing Systems 15, MIT Press.

Hyvärinen, A. (1999). Fast and robust fixed-point algorithms for independent component analysis. IEEE Transactions on Neural Networks, 10(3), 626-634. doi: 10.1109/72.761722.

Hyvärinen, A., & Hoyer, P. O. (2001). A two-layer sparse coding model learns simple and complex cell receptive fields and topography from natural images. Vision Research, 41(18), 2413-23.

Hyvärinen, A., & Oja, E. (1997). A fast fixed-point algorithm for independent component analysis. Neural Comput., 9(7), 1483-1492.

Johnson, M. H., & Morton, J. (1991). Biology and cognitive development: The case of face recognition. Blackwell Oxford.

Jones, J. P., & Palmer, L. A. (1987). An evaluation of the two-dimensional Gabor filter model of simple receptive fields in cat striate cortex. J Neurophysiol, 58(6), 1233-1258.

Kalb, R. G., & Hockfield, S. (1992). Activity-dependent development of spinal cord motor neurons. Brain Research Reviews, 17(3), 283-289.

Kanwisher, N., McDermott, J., & Chun, M. M. (1997). The Fusiform Face Area: A Module in Human Extrastriate Cortex Specialized for Face Perception. J. Neurosci., 17(11), 4302-4311.

Kohn, A., & Smith, M. A. (2005). Stimulus Dependence of Neuronal Correlation in Primary Visual Cortex of the Macaque. J. Neurosci., 25(14), 3661-3673. doi: 10.1523/JNEUROSCI.5106-04.2005.

Krug, K., Akerman, C. J., & Thompson, I. D. (2001). Responses of Neurons in Neonatal Cortex and Thalamus to Patterned Visual Stimulation Through the Naturally Closed Lids. J Neurophysiol, 85(4), 1436-1443.

Laughlin, S. B., de Ruyter van Steveninck, R. R., & Anderson, J. C. (1998). The metabolic cost of neural information. Nat Neurosci, 1(1), 36-41. doi: 10.1038/236.

Leake, P. A., Hradek, G. T., Chair, L., & Snyder, R. L. (2006). Neonatal Deafness Results in Degraded Topographic Specificity (Frequency Resolution) of Auditory Nerve Projections to the Cochlear Nucleus in Cats. The Journal of comparative neurology, 497(1), 13–31. doi: 10.1002/cne.20968.

Lehky, S. R., Sejnowski, T. J., & Desimone, R. (2005). Selectivity and sparseness in the responses of striate complex cells. Vision Research, 45(1), 57-73. doi: 10.1016/j.visres.2004.07.021.

Lennie, P. (2003). The cost of cortical computation. Current Biology: CB, 13(6), 493-497.

Li, Y., Fitzpatrick, D., & White, L. E. (2006). The development of direction selectivity in ferret visual cortex requires early visual experience. Nat Neurosci, 9(5), 676-681. doi: 10.1038/nn1684.

Linsker, R. (1986). From basic network principles to neural architecture: emergence of orientation columns. Proceedings of the National Academy of Sciences of the United States of America, 83(22), 8779-8783.

Lippe, W. R. (1994). Rhythmic spontaneous activity in the developing avian auditory system. J. Neuroscience, 14, 1486-95.

Lisman, J. E. (1997). Bursts as a unit of neural information: making unreliable synapses reliable. Trends in Neurosciences, 20(1), 38-43.

Lobue, V., & DeLoache, J. S. (2008). Detecting the snake in the grass: attention to fear-relevant stimuli by adults and young children. Psychological Science, 19(3), 284-289. doi: 10.1111/j.1467-9280.2008.02081.x.

Lowel, S., & Singer, W. (1992). Selection of intrinsic horizontal connections in the visual cortex by correlated neuronal activity. Science, 255(5041), 209-212. doi: 10.1126/science.1372754.

Maffei, L., & Galli-Resta, L. (1990). Correlation in the discharges of neighboring rat retinal ganglion cells during prenatal life. Proceedings of the National Academy of Sciences of the United States of America, 87(7), 2861-2864.

Malarz, K., & Galam, S. (2005). Square-lattice site percolation at increasing ranges of neighbor bonds. Physical Review E, 71(1), 016125. doi: 10.1103/PhysRevE.71.016125.

Marcelja, S. (1980). Mathematical description of the responses of simple cortical cells. Journal of the Optical Society of America, 70(11), 1297-1300.

Marr, D. (1982). Vision. Freeman San Francisco.

Maruko, I., Zhang, B., Tao, X., Tong, J., Smith, E. L., & Chino, Y. M. (2008). Postnatal Development of Disparity Sensitivity in Visual Area 2 (V2) of Macaque Monkeys. J Neurophysiol, 100(5), 2486-2495. doi: 10.1152/jn.90397.2008.

Meister, M., Wong, R. O., Baylor, D. A., & Shatz, C. J. (1991). Synchronous bursts of action potentials in ganglion cells of the developing mammalian retina. Science, 252(5008), 939-943.

Miikkulainen, R., Bednar, J. A., Choe, Y., & Sirosh, J. (2005). Computational Maps in

the Visual Cortex (1st ed.). Springer.

Miller, K. D., Keller, J. B., & Stryker, M. P. (1989). Ocular dominance column development: analysis and simulation. Science, 245(4918), 605-615.

Miller, K. D. (1994). A model for the development of simple cell receptive fields and the ordered arrangement of orientation columns through activity-dependent competition between ON- and OFF-center inputs. The Journal of Neuroscience, 14, 409–441.

Mineka, S., Davidson, M., Cook, M., & Keir, R. (1984). Observational conditioning of snake fear in rhesus monkeys. Journal of Abnormal Psychology, 93(4), 355-372.

Murphy, K. M., Duffy, K. R., Jones, D. G., & Mitchell, D. E. (2001). Development of cytochrome oxidase blobs in visual cortex of normal and visually deprived cats. Cerebral Cortex, 11(2), 122-135.

O'Donovan, M. J., Chub, N., & Wenner, P. (1998). Mechanisms of spontaneous activity in developing spinal networks. Journal of Neurobiology, 37(1), 131-145.

O'Donovan, M. J., Wenner, P., Chub, N., Tabak, J., & Rinzel, J. (1998). Mechanisms of spontaneous activity in the developing spinal cord and their relevance to locomotion. Annals of the New York Academy of Sciences, 860, 130-41.

O'Kusky, J., & Colonnier, M. (1982). A laminar analysis of the number of neurons, glia, and synapses in the adult cortex (area 17) of adult macaque monkeys. The Journal of Comparative Neurology, 210(3), 278-290. doi: 10.1002/cne.902100307.

O'Leary, D. D., Ruff, N. L., & Dyck, R. H. (1994). Development, critical period plasticity, and adult reorganizations of mammalian somatosensory systems. Current Opinion in Neurobiology, 4(4), 535-544.

O'Leary, D. D. M., Ruff, N. L., & Dyck, R. H. (1998). Development, critical period plasticity, and adult reorganizations of mammalian somatosensory systems. Findings and Current Opinion in Cognitive Neuroscience, 4, 307.

Ohzawa, I., & Freeman, R. D. (1986a). The binocular organization of simple cells in the cat's visual cortex. Journal of Neurophysiology, 56(1), 221-242.

Ohzawa, I., & Freeman, R. D. (1986b). The binocular organization of complex cells in the cat's visual cortex. Journal of Neurophysiology, 56(1), 243-259.

Olshausen, B. A., & Field, D. J. (1996). Emergence of simple-cell receptive field properties by learning a sparse code for natural images. Nature, 381(6583), 607-9. doi: 10.1038/381607a0.

Olshausen, B. A., & Field, D. J. (1997). Sparse coding with an overcomplete basis set:

a strategy employed by V1? Vision Research, 37(23), 3311-3325.

Ozaki, T., & Kaplan, E. (2006). Brainstem input modulates globally the transmission through the lateral geniculate nucleus. The International Journal of Neuroscience, 116(3), 247-264. doi: 10.1080/00207450500403033.

Pollen, D., & Ronner, S. (1983). Visual cortical neurons as localized spatial frequency filters. IEEE Transactions on Systems, Man and Cybernetics, 13(5), 907-916.

Provine, R. R. (1972). Ontogeny of bioelectric activity in the spinal cord of the chick embryo and its behavioral implications. Brain Research, 41(2), 365-78.

Qian, N. (1997). Binocular disparity and the perception of depth. Neuron, 18(3), 359-368.

Qian, N., & Zhu, Y. (1997). Physiological computation of binocular disparity. Vision Research, 37(13), 1811-1827. doi: 10.1016/S0042-6989(96)00331-8.

Ringach, D. L. (2004). Haphazard wiring of simple receptive fields and orientation columns in visual cortex. Journal of Neurophysiology, 92(1), 468-476. doi: 10.1152/jn.01202.2003.

Ringach, D. L. (2007). On the origin of the functional architecture of the cortex. PLoS ONE, 2(2), e251. doi: 10.1371/journal.pone.0000251.

Rozell, C. J., Johnson, D. H., Baraniuk, R. G., & Olshausen, B. A. (2008). Sparse coding via thresholding and local competition in neural circuits. Neural Comput., 20(10), 2526-2563.

Schneidman, E., Berry, M. J., Segev, R., & Bialek, W. (2006). Weak pairwise correlations imply strongly correlated network states in a neural population. Nature, 440(7087), 1007-1012. doi: 10.1038/nature04701.

Schroeder, M. (1992). Fractals, Chaos, Power Laws: Minutes from an Infinite Paradise. W. H. Freeman.

Schwartz, T. H., Rabinowitz, D., Unni, V., Kumar, V. S., Smetters, D. K., Tsiola, A., et al. (1998). Networks of Coactive Neurons in Developing Layer 1. Neuron, 20(3), 541-552. doi: 10.1016/S0896-6273(00)80993-9.

Scoville, W. B., & Milner, B. (1957). Loss of recent memory after bilateral hippocampal lesions. Journal of Neurology, Neurosurgery, and Psychiatry, 20(1), 11-21.

Sengpiel, F. (2006). Motion perception is learned, not innate. Nat Neurosci, 9(5), 591-592. doi: 10.1038/nn0506-591.

Sengpiel, F., & Kind, P. C. (2002). The role of activity in development of the visual system. Current Biology: CB, 12(23), 818-826.

Sengpiel, F., Stawinski, P., & Bonhoeffer, T. (1999). Influence of experience on orientation maps in cat visual cortex. Nat Neurosci, 2(8), 727-732. doi: 10.1038/11192.

Sernagor, E., & Grzywacz, N. M. (1996). Influence of spontaneous activity and visual experience on developing retinal receptive fields. Current Biology, 6(11), 1503-1508.

Shatz, C., & Stryker, M. (1988). Prenatal tetrodotoxin infusion blocks segregation of retinogeniculate afferents. Science, 242(4875), 87-89. doi: 10.1126/science.3175636.

Shouval, H., Intratorj, N., & Cooper, L. N. (1997). BCM network develops orientation selectivity and ocular dominance in natural scene environment. Vision Research, 37(23), 3339-3342. doi: 10.1016/S0042-6989(97)00087-4.

Skaggs, W. E., & McNaughton, B. L. (1996). Replay of neuronal firing sequences in rat hippocampus during sleep following spatial experience. Science, 271(5257), 1870-1873.

Slater, A. (1993). Visual perceptual abilities at birth: Implications for face perception. In B. de Boysson-Bardies (Ed.), Developmental neurocognition: Speech and face processing in the first year of life (pp. 125-134). Boston: Kluwer.

Squire, L. R., & Alvarez, P. (1995). Retrograde amnesia and memory consolidation: a neurobiological perspective. Current Opinion in Neurobiology, 5(2), 169-177. doi: 10.1016/0959-4388(95)80023-9.

Stickgold, R. (2005). Sleep-dependent memory consolidation. Nature, 437(7063), 1272-1278. doi: 10.1038/nature04286.

Strata, F., Atzori, M., Molnar, M., Ugolini, G., Tempia, F., & Cherubini, E. (1997). A Pacemaker Current in Dye-Coupled Hilar Interneurons Contributes to the Generation of Giant GABAergic Potentials in Developing Hippocampus. J. Neurosci., 17(4), 1435-1446.

Stryker, M. P., & Harris, W. A. (1986). Binocular impulse blockade prevents the formation of ocular dominance columns in cat visual cortex. J. Neuroscience, 6(8), 2117-2133.

Stryker, M. P., & Strickland, S. L. (1984). Physiological segregation of ocular dominance columns depends on the pattern of afferent electrical activity. Invest Opthalmol Vis Sci, 25, 278.

Swindale, N. V. (1996). The development of topography in the visual cortex: a review of models. Network, 7(2), 161-247. doi: 10.1088/0954-898X/7/2/002.

Tanaka, S., Miyashita, M., & Ribot, J. (2004). Roles of visual experience and intrinsic mechanism in the activity-dependent self-organization of orientation maps: theory and experiment. Neural Netw., 17(8-9), 1363-1375.

Todorov, E. (2004). Optimality principles in sensorimotor control. Nature Neuroscience, 7(9), 907-915. doi: 10.1038/nn1309.

Tritsch, N. X., Yi, E., Gale, J. E., Glowatzki, E., & Bergles, D. E. (2007). The origin of spontaneous activity in the developing auditory system. Nature, 450(7166), 50-55. doi: 10.1038/nature06233.

Tyler, C. W., & Clarke, M. B. (1990). Autostereogram. Stereoscopic Displays and Applications, 1256, 182-197.

Uesaka, N., Ruthazer, E. S., & Yamamoto, N. (2006). The Role of Neural Activity in Cortical Axon Branching. Neuroscientist, 12(2), 102-106.

Utgoff, P. E., & Mitchell, T. M. (1982). Acquisition of Appropriate Bias for Inductive Concept Learning (pp. 414-417). Menlo Park, CA: AAAI Press.

Valsalam, V., Bednar, J., & Miikkulainen, R. (2007). Developing Complex Systems Using Evolved Pattern Generators. Evolutionary Computation, IEEE Transactions on, 11(2), 181-198. doi: 10.1109/TEVC.2006.890272.

van Hateren, J. H., & van der Schaaf, A. (1998). Independent component filters of natural images compared with simple cells in primary visual cortex. Proceedings of the Royal Society B: Biological Sciences, 265(1394), 359–366.

van Hateren, J. H., & Snippe, H. P. (2006). Phototransduction in primate cones and blowfly photoreceptors: different mechanisms, different algorithms, similar response. Journal of Comparative Physiology. A, Neuroethology, Sensory, Neural, and Behavioral Physiology, 192(2), 187-97. doi: 10.1007/s00359-005-0060-y.

Victor, J. D., Mechler, F., Repucci, M. A., Purpura, K. P., & Sharpee, T. (2006). Responses of V1 Neurons to Two-Dimensional Hermite Functions. J Neurophysiol, 95(1), 379-400. doi: 10.1152/jn.00498.2005.

von der Malsburg, C. (1973). Self-organization of orientation sensitive cells in the striate cortex. Kybernetik, 14(2), 85-100.

Walsh, E. J., McGee, J., & Javel, E. (1986). Development of auditory-evoked potentials in the cat. I. Onset of response and development of sensitivity. The Journal of the Acoustical Society of America, 79(3), 712-724. doi: 10.1121/1.393461.

Warland, D. K., Huberman, A. D., & Chalupa, L. M. (2006). Dynamics of Spontaneous Activity in the Fetal Macaque Retina during Development of Retinogeniculate Pathways. J. Neurosci., 26(19), 5190-5197.

Webber, A. L., & Wood, J. (2005). Amblyopia: prevalence, natural history, functional effects and treatment. Clinical & Experimental Optometry: Journal of the Australian Optometrical Association, 88(6), 365-75.

Weber, C. (2001). Self-Organization of Orientation Maps, Lateral Connections, and Dynamic Receptive Fields in the Primary Visual Cortex. In Artificial Neural Networks — ICANN 2001 (pp. 1147-1152).

Weliky, M., & Katz, L. C. (1997). Disruption of orientation tuning in visual cortex by artificially correlated neuronal activity. Nature, 386(6626), 680-5. doi: 10.1038/386680a0.

Weliky, M., & Katz, L. C. (1999). Correlational structure of spontaneous neuronal activity in the developing lateral geniculate nucleus in vivo. Science, 285(5427), 599-604.

White, L. E., Coppola, D. M., & Fitzpatrick, D. (2001). The contribution of sensory experience to the maturation of orientation selectivity in ferret visual cortex. Nature, 411(6841), 1049-1052. doi: 10.1038/35082568.

Willmore, B., & Tolhurst, D. J. (2001). Characterizing the sparseness of neural codes. Network, 12(3), 255-270.

Wong, R. O. (1999). Retinal waves and visual system development. Annual Review of Neuroscience, 22, 29-47. doi: 10.1146/annurev.neuro.22.1.29.

Wong, R. O. L., & Ghosh, A. (2002). Activity-dependent regulation of dendritic growth and patterning. Nat Rev Neurosci, 3(10), 803-812. doi: 10.1038/nrn941.

Wong, W. T., Sanes, J. R., & Wong, R. O. (1998). Developmentally regulated spontaneous activity in the embryonic chick retina. J. Neuroscience, 18(21), 8839-8852.

Yu, C. R., Power, J., Barnea, G., O'Donnell, S., Brown, H. E. V., Osborne, J., et al. (2004). Spontaneous neural activity is required for the establishment and maintenance of the olfactory sensory map. Neuron, 42(4), 553-566.

Yuste, R., Nelson, D. A., Rubin, W. W., & Katz, L. C. (1995). Neuronal domains in developing neocortex: mechanisms of coactivation. Neuron, 14(1), 7-17.

Zhou, Z. J. (1998). Direct Participation of Starburst Amacrine Cells in Spontaneous Rhythmic Activities in the Developing Mammalian Retina. J. Neurosci., 18(11), 4155-4165.

Zito, T., Wilbert, N., Wiskott, L., & Berkes, P. (2008). Modular Toolkit for Data Processing (MDP): A Python Data Processing Framework. Frontiers in neuroinformatics, 2.